Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática

Apunte del Curso

Cálculo en Varias Variables

Autores:
Patricio Felmer
Alejandro Jofré

Con la colaboración de:
Paul Bosch
Matías Bulnes
Nicolás Hernández
Arturo Prat
Luis Rademacher
Mauricio Vargas
José Zamora

13 de junio de 2013

Índice general

Índice de figuras

Introducción

En ciencias e ingeniería se emplean numerosos modelos para describir matemáticamente fenómenos de diferente índole que van desde el cálculo de estructuras hasta fenómenos económicos sociales pasando por la mecánica de fluidos, transferencia de calor, equilibrios químicos, planificación y gestión de procesos, biotecnología, astronomía, física del estado sólido, materiales, minería, sólo por mencionar algunas que se cultivan en la facultad. Para esta tarea el cálculo de una variable muchas veces es insuficiente, pues la realidad incorpora múltiples variables y sus interacciones para el estudio de estos fenómenos.

Los autores.

CAPÍTULO 1

Cálculo Diferencial en $\mathbb{R}^n$

Nos interesa ampliar las herramientas aprendidas en el curso de Cálculo a funciones de varias variables. Para esto debemos introducir, entre otros, los conceptos de derivada parcial y diferencial. Con ambos conceptos podremos entender diversas herramientas que nos permitirán continuar la tarea de maximizar o minimizar funciones que pueden estar sujetas a una o más restricciones.

1.1. Repaso de algunos teoremas de Cálculo

Los siguientes teoremas, vistos en el curso de Cálculo serán de utilidad en todo lo que sigue. Cuando trabajemos algunos teoremas en $\mathbb{R}^n$ nos podremos dar cuenta que sus demostraciones son análogas a las del curso de Cálculo.

Teorema 1.1. *Sea $[a, b]$ un intervalo cerrado y acotado y $f : [a, b] \to \mathbb{R}$ continua tal que $f(a)$ y $f(b)$ tienen distinto signo, entonces existe $c \in [a, b]$ tal que $f(c) = 0$.*

Demostración. Sin pérdida de generalidad supongamos que $f(a) < 0$ y $f(b) > 0$. Consideremos el punto medio del intervalo $c = \frac{a+b}{2}$, para el cual existen tres posibilidades:

1. $f(c) < 0$: en este caso nos restringiremos al intervalo $[a_1, b_1]$ con $a_1 = c$ y $b_1 = b$.
2. $f(c) = 0$: en este caso concluye la demostración.
3. $f(c) > 0$: en este caso nos restringiremos al intervalo $[a_1, b_1]$ con $a_1 = a$ y $b_1 = c$.

Si realizamos de forma consecutiva el proceso de anterior, y si siempre nos mantenemos dentro de los casos 1 y 3, habremos generado una sucesión de intervalos $[a_n, b_n] \subset [a_{n-1}, b_{n-1}] \subset \ldots \subset [a, b]$ tales que $f(a_n) < 0$, $f(b_n) > 0$ y $b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2}$ para todo n. De esta forma, $b_n - a_n = \frac{b-a}{2^n}$ y para $n \to \infty$ se tendrá que $a_n - b_n \to 0$ y $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Sea $e = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$, claramente $e \in (a, b)$ y además, como f continua:

$$f(e) = f\left(\lim_{n\to\infty} a_n\right) = \lim_{n\to\infty} f(a_n) \leq 0$$

donde la desigualdad se tiene ya que $f(a_n) < 0$ para todo n. Análogamente se cumple:

$$f(e) = f\left(\lim_{n\to\infty} b_n\right) = \lim_{n\to\infty} f(b_n) \geq 0$$

pues $f(b_n) < 0$ para todo n. Se concluye entonces que $f(e) = 0$. ■

Teorema 1.2. (Teorema del valor intermedio en $\mathbb{R}$)
Sea $[a, b]$ un intervalo cerrado y acotado y $f : [a, b] \to \mathbb{R}$ continua. Si $f(a) \neq f(b)$, entonces para todo número k comprendido entre $f(a)$ y $f(b)$ existe $c \in (a, b)$ tal que $f(c) = k$.

Demostración. Sin pérdida de generalidad supongamos que $f(a) < f(b)$. Definamos $g(x) = f(x) - k$ y entonces $g(a) = f(a) - k < 0$ y $g(b) = f(b) - k > 0$. De acuerdo al teorema 1.1 existe $c \in (a, b)$ tal que $g(c) = 0$, lo que es equivalente a $f(c) = k$. ■

Teorema 1.3. (Teorema de Bolzano-Weierstrass en $\mathbb{R}$)
Toda sucesión acotada en $\mathbb{R}$ tiene al menos una subsucesión convergente.

Demostración. Sea $\{x_n\}$ una sucesión acotada en $\mathbb{R}$. Entonces existen $a_0, b_0 \in \mathbb{R}$ tales que:

$$\forall n \in \mathbb{N}, a_0 \leq x_n \leq b_n$$

Consideremos ahora $c_0 = \frac{a_0+b_0}{2}$. En al menos uno de los intervalos $[a_0, c_0]$ y $[c_0, b_0]$ hay infinitos términos de $\{x_n\}$, llamemos $[a_1, b_1]$ a dicho intervalo y hagamos lo mismo consecutivamente de forma de obtener una sucesión de intervalos $[a_n, b_n] \subset [a_{n-1}, b_{n-1}] \subset \ldots \subset [a, b]$ tales que $[a_n, b_n]$ contiene infintos términos de $\{x_n\}$ y $b_n - a_n = \frac{b_{n-1}-a_{n-1}}{2}$ para todo n. Consideremos ahora la subsucesión $\{x_{f(n)}\}$ donde los índices $f(n)$ están definidos por:

$$f(1) = \text{mín}\{i \in \mathbb{N} \mid x_i \in [a_1, b_1]\}$$

$$f(n+1) = \text{mín}\{i > f(n) \mid x_i \in [a_{n+1}, b_{n+1}]\}$$

Se tiene entonces:

$$\forall n \in \mathbb{N}, a_n \leq x_{f(n)} \leq b_n$$

Notemos que la sucesión $\{a_n\}$ es monótona creciente y por lo tanto converge a un real a. La sucesión $\{b_n\}$ en cambio, es monótona decreciente y converge a un real b. Notando además que $b_n - a_n = \frac{b-a}{2^n}$ para todo n, se tiene que $\text{lím}_{n\to\infty} a_n = \text{lím}_{n\to\infty} b_n$, es decir $a = b$. Por el teorema del sandwich concluímos que $\text{lím}_{n\to\infty} x_{f(n)} = a = b$. ■

Teorema 1.4. *Sea $[a, b]$ un intervalo cerrado y acotado y $f : [a, b] \to \mathbb{R}$ continua. Entonces f es acotada y alcanza su mínimo y su máximo en $[a, b]$.*

Demostración. Probaremos el resultado para el máximo y la parte del mínimo quedará propuesta (es análoga). Sea $M = \sup\{f(x) : a \leq x \leq b\}$, eventualmente $M = \infty$, y consideremos una sucesión $\{x_n\}$ tal que $f(x_n)$ converge a M. Por el teorema 1.3 existe una subsucesión x_{n_k} que converge a un punto $\bar{x} \in [a, b]$. Como f es continua se tiene que la subsucesión $f(x_{n_k})$ converge a $f(\bar{x})$ y por lo tanto $f(\bar{x}) = M$. Esto prueba simultáneamente que M es finito y que el supremo (máximo) se alcanza. ■

Teorema 1.5. *Sea $[a, b]$ un intervalo cerrado y acotado y $f : [a, b] \to \mathbb{R}$ continua y diferenciable en (a, b). Si f alcanza un máximo (o mínimo) en algún punto $c \in (a, b)$ entonces $f'(c) = 0$.*

Demostración. Haremos la demostración para el caso del máximo. La demostración para el caso del mínimo es análoga y queda de propuesta. Si f tiene un máximo en c entonces:

$$f(c + h) \leq f(c) \ \forall h \text{ tal que } c + h \in [a, b]$$

De esta forma, $f(c+h) - f(c) \leq 0$. Así para todo $h > 0$ pequeño se tiene que:

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

y esto implica que $f'(c) \leq 0$. Por otro lado, para todo $h < 0$ pequeño se tiene que:

$$\frac{f(c+h) - f(c)}{h} \geq 0$$

lo que implica que $f'(c) \geq 0$ y por lo tanto se concluye que $f'(c) = 0$. ∎

Teorema 1.6. (Teorema de Rolle en $\mathbb{R}$)
Sean $[a,b]$ cerrado y acotado y $f : [a,b] \to \mathbb{R}$ continua y diferenciable. Si $f(a) = f(b)$, entonces existe al menos un $c \in (a,b)$ tal que $f'(c) = 0$.

Demostración. Tenemos tres casos posibles:

1. Si $f(c) < f(a)$ para algún $c \in (a,b)$. Por el teorema 1.4 sabemos que existe $c \in [a,b]$ donde f alcanza su mínimo y en este caso necesariamente $c \in (a,b)$. De acuerdo al teorema 1.5 $f'(c) = 0$.
2. Si $f(a) = f(c)$ $\forall c \in (a,b)$. Entonces, por ser f constante, su derivada es nula en (a,b) y se cumple el teorema.
3. Si $f(c) > f(a)$ para algún $c \in (a,b)$. Por el teorema 1.4 sabemos que existe $c \in [a,b]$ donde f alcanza su máximo y en este caso necesariamente $c \in (a,b)$. De acuerdo al teorema 1.5 $f'(c) = 0$.

∎

Teorema 1.7. (Teorema del valor medio en $\mathbb{R}$)
Sean $[a,b]$ un intervalo cerrado y acotado y $f : [a,b] \to \mathbb{R}$ continua y diferenciable en (a,b). Entonces, existe un punto $c \in (a,b)$ tal que:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Demostración. Definamos

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

Se tiene que g es continua en $[a,b]$ y diferenciable en (a,b). Además, observemos que

$$g(a) = f(a) \qquad g(b) = f(a)$$

lo que implica $g(a) = g(b)$ por lo tanto podemos aplicar el teorema 1.6. Entonces, existe $c \in (a,b)$ tal que $g'(c) = 0$ lo que es equivalente a:

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

∎

Teorema 1.8. (Primer teorema fundamental del cálculo)
Sean $f : [a, b] \to \mathbb{R}$ *continua y* $c \in [a, b]$*, entonces la función* F *definida por*

$$F(x) = \int_a^x f(x)dx$$

es derivable en (a, b) *y además* $F'(x) = f(x)$ *en* (a, b)*.*

Demostración. Sea $c \in (a, b)$. Debemos demostrar que el límite

$$F'(c) = \lim_{h \to 0} \frac{F(c+h) - F(c)}{h}$$

existe y vale $f(c)$. Notemos que

$$F(c+h) - F(c) = \int_a^{c+h} f(x)dx - \int_a^c f(x)dx = \int_c^{c+h} f(x)dx$$

Consideremos por separado los casos $h > 0$ y $h < 0$:

1. Sea $h > 0$: Como f es continua en $[c, c+h]$, se tiene que existen valores a_1 y b_1 en $[c, c+h]$ tales que
$$f(a_1) \leq f(x) \leq f(b_1) \qquad \forall x \in [c, c+h]$$
Integrando en $[c, c+h]$
$$f(a_1)h \leq F(c+h) - F(c) \leq f(b_1)h$$
$$f(a_1) \leq \frac{F(c+h) - F(c)}{h} \leq f(b_1)$$
Si $h \to 0^+$ entonces $a_1 \to c$ y $b_1 \to c$. Como f es continua $f(a_1) \to f(c)$ y $f(b_1) \to f(c)$. Entonces,
$$\lim_{h \to 0^+} \frac{F(c+h) - F(c)}{h} = f(c) \qquad (*)$$

2. Sea $h < 0$: Como f es continua en $[c+h, c]$, se tiene que existen valores a_2 y b_2 en $[c+h, c]$ tales que
$$f(a_2) \leq f(x) \leq f(b_2) \qquad \forall x \in [c+h, c]$$
Integrando en $[c+h, c]$
$$f(a_2)(-h) \leq -(F(c+h) - F(c)) \leq f(b_2)(-h)$$
$$f(a_2) \leq \frac{F(c+h) - F(c)}{h} \leq f(b_2)$$
Si $h \to 0^-$ entonces $a_2 \to c$ y $b_2 \to c$. Como f es continua, $f(a_2) \to f(c)$ y $f(b_2) \to f(c)$. Entonces,
$$\lim_{h \to 0^-} \frac{F(c+h) - F(c)}{h} = f(c) \qquad (**)$$

De (*) y (**) se obtiene

$$\lim_{h\to 0^-} \frac{F(c+h)-F(c)}{h} = \lim_{h\to 0^+} \frac{F(c+h)-F(c)}{h} = f(c)$$

Por lo tanto $F'(c) = f(c)$. ■

Teorema 1.9. (Segundo teorema fundamental del cálculo)
Sea $f : [a,b] \to \mathbb{R}$ integrable. Si existe una función $F : [a,b] \to \mathbb{R}$ continua y derivable tal que $F'(x) = f(x)$ en (a,b), entonces

$$\int_a^b f(x)dx = F(b) - F(a)$$

Demostración. Sea $P = \{x_0, \dots, x_n\}$ una partición cualquiera del intervalo $[a,b]$, entonces en cada intervalo $[x_{i-1}, x_i]$ la función $F(x)$ cumple las hipótesis del teorema del valor medio (teorema 1.7), es decir

$$F(x_i) - F(x_{i-1}) = F'(c_i)(x_i - x_{i-1})$$

Como $F'(x) = f(x)\ \forall x \in [a,b]$, entonces $F'(c_i) = f(c_i)$ y además

$$m_i(f)(x_i - x_{i-1}) \leq f(c_i)(x_i - x_{i-1}) \leq M_i(f)(x_i - x_{i-1})$$

Lo que es equivalente a

$$m_i(f)(x_i - x_{i-1}) \leq F(x_i) - F(x_{i-1}) \leq M_i(f)(x_i - x_{i-1})$$

Aplicando $\sum_{i=1}^n (\cdot)$ se obtiene

$$s(f,P) \leq F(b) - F(a) \leq S(f,P)$$

Luego, como lo anterior es válido para todas las particiones de $[a,b]$ y f es integrable en $[a,b]$

$$\int_a^b f(x)dx = F(b) - F(a)$$

■

Teorema 1.10. (Teorema del valor medio en $\mathbb{R}$, versión integral)
Sean $[a,b]$ un intervalo cerrado y acotado y $f : [a,b] \to \mathbb{R}$ continua. Entonces, existe un punto $c \in (a,b)$ tal que:

$$\int_a^b f(s)ds = f(c)$$

Demostración. Definamos $F : [a,b] \to \mathbb{R}$ como $F(x) = \int_a^x f(s)ds$. Entonces por el teorema 1.8 F es continua, diferenciable en (a,b) y $F'(x) = f(x)$. Usando primero el teorema 1.9 y luego el teorema 1.7, se tiene que

$$\int_a^b f(x)dx = F(b) - F(a) = F'(c) = f(c)$$

para algún $c \in (a,b)$. ■

1.2. Base algebraica y geométrica de $\mathbb{R}^n$

Definición 1.1. Dotamos al conjunto $\mathbb{R}^n$ de una estructura de espacio vectorial mediante las siguientes operaciones: para $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, y $\lambda \in \mathbb{R}$ definimos

Suma:

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \ldots, x_n + y_n)$$

Producto por escalar:

$$\lambda \boldsymbol{x} = (\lambda x_1, \ldots, \lambda x_n)$$

El espacio vectorial $\mathbb{R}^n$ tiene dimensión n. Entre las muchas posibles bases de $\mathbb{R}^n$ nos interesará considerar, por su simplicidad, la llamada base canónica $\{\boldsymbol{e}_i\}_{i=1}^n$, donde $\boldsymbol{e}_i = (0, ..., 1, ..., 0)$ con el uno en la posición i. Así todo $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ se puede representar en términos de la base canónica como

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i$$

Habiendo ya definido la estructura algebraica de $\mathbb{R}^n$ vamos a introducir la estructura geométrica de $\mathbb{R}^n$ a través del producto interno (o producto punto)

Definición 1.2. Dados $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, se define el producto interno o punto de $\boldsymbol{x}$ e $\boldsymbol{y}$ como

$$\boldsymbol{x} \cdot \boldsymbol{y} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} x_i y_i$$

La siguiente proposición resume las propiedades básicas del producto interno. Su demostración es muy simple.

Proposición 1.1. (Propiedades del producto interno)

1. Positividad: Para todo $\boldsymbol{x} \in \mathbb{R}^n$ se tiene $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ y $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ sí y sólo sí $\boldsymbol{x} = 0$.
2. Linealidad: Para todo $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$ $\langle \lambda \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
3. Simetría: Para todo $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$.

La noción de producto interno induce de manera natural la noción de norma o longitud de un vector.

Definición 1.3. Se define la norma de un vector $\boldsymbol{x} \in \mathbb{R}^n$ como

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}$$

La siguiente proposición establece una desigualdad entre producto interno y norma de vectores de $\mathbb{R}^n$. Ella nos permite definir la noción de ángulo entre vectores de $\mathbb{R}^n$, dejando en evidencia que el producto interno determina la geometría de $\mathbb{R}^n$.

Proposición 1.2. (Desigualdad de Cauchy-Schwarz)

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\| \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

Además, se cumple la igualdad si y sólo si $\boldsymbol{x}$ es múltiplo escalar de $\boldsymbol{y}$ o uno de ellos es cero.

Demostración. Consideremos $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ cualesquiera que mantendremos fijos y sea $t \in \mathbb{R}$. Entonces por las propiedades del producto interno tenemos que

$$0 \leq \langle \boldsymbol{x} + t\boldsymbol{y}, \boldsymbol{x} + t\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{x} \rangle + 2t\langle \boldsymbol{x}, \boldsymbol{y} \rangle + t^2 \langle \boldsymbol{y}, \boldsymbol{y} \rangle$$

Si $\boldsymbol{y} = \mathbf{0}$ entonces la desigualdad que se desea probar naturalmente vale. Si $\boldsymbol{y} \neq \mathbf{0}$ entonces notamos que la expresión de arriba determina una función cuadrática en t que puede anularse a lo más una vez. Esto implica que el discriminante debe ser negativo o nulo, es decir,

$$4\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 - 4\langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle \leq 0$$

de donde se obtiene la desigualdad deseada. Para terminar, cuando $\boldsymbol{x}$ es múltiplo de $\boldsymbol{y}$ entonces claramente se tiene la igualdad. Queda propuesto probar la recíproca. ■

La desigualdad de Cauchy-Schwarz nos permite definir la noción de ángulo entre vectores.

Definición 1.4. Dados $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ llamaremos ángulo entre $\boldsymbol{x}$ e $\boldsymbol{y}$ a:

$$\theta = \arccos \left(\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\| \|\boldsymbol{y}\|} \right)$$

entendemos que $\theta \in [0, \pi]$.

Con esta definición podemos hablar de vectores ortogonales cuando el ángulo entre ellos es de 90^o, es decir, cuando $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.

Además, también se obtiene la validez del teorema del coseno:

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| \cos(\theta)$$

Cuando los vectores son ortogonales tenemos el teorema de Pitágoras.

Como ya dijimos, el producto interno induce la noción de norma, la que le da a $\mathbb{R}^n$ su carácter topológico, como ya veremos. Por el momento veamos las propiedades básicas de la norma.

Proposición 1.3. (Propiedades de la norma)

1. Positividad: Para todo $\boldsymbol{x} \in \mathbb{R}^n$ $\|\boldsymbol{x}\| \geq 0$ y $\|\boldsymbol{x}\| = 0$ si y sólo si $\boldsymbol{x} = 0$.
2. Homogeneidad: Para todo $\boldsymbol{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}$ $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$.
3. Desigualdad triangular: Para todo $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

Demostración. 1. y 2. son directas del las propiedades del producto interno y la definición de norma.

La Desigualdad Triangular es una consecuencia de la desigualdad de Cauchy-Schwarz. Para $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ tenemos

$$\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle = \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 \leq \|\boldsymbol{x}\|^2 + 2\|\boldsymbol{x}\|\|\boldsymbol{y}\| + \|\boldsymbol{y}\|^2$$

de donde se obtiene

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 \leq (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2$$

■

Nota 1.1. De ahora en adelante preferimos denotar el producto punto entre vectores $\boldsymbol{x}$ e $\boldsymbol{y}$ como $\boldsymbol{x} \cdot \boldsymbol{y}$.

1.3. Funciones con valores en $\mathbb{R}^m$

Definición 1.5. Llamaremos a f función a valores en $\mathbb{R}^m$ si $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$.

Observamos que el argumento de f es un vector $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y que la imagen de $\boldsymbol{x}$ es un vector de $\mathbb{R}^m$. Así $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x}))$, donde las funciones $f_i : D \subseteq \mathbb{R} \to \mathbb{R}$, para cada i, se conocen como funciones coordenadas.

Para el estudio de las funciones de $\mathbb{R}^n$ a valores en $\mathbb{R}^m$ vamos a desarrollar las herramientas del Cálculo Diferencial. Sin embargo, la posibilidad de dibujar en el caso de dimensiones pequeñas, es siempre algo muy útil. Más aún ahora que tenemos programas computacionales (Matlab, Wolfram Mathematica, Gnu Octave, etc.) muy eficientes para esta tarea. A continuación damos alguna terminología.

Definición 1.6. Llamaremos grafo de una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ al conjunto:

$$G(f) = \{(\boldsymbol{x}, f(\boldsymbol{x})) : \boldsymbol{x} \in D\}$$

Notemos que $G(f) \subset \mathbb{R}^{n+m}$. Este se podrá dibujar cuando $m = 1$ y $n = 1$ o $n = 2$. En el primer caso el grafo es una curva y en el segundo una superficie.

Definición 1.7. Si $m = 1$, dado $c \in \mathbb{R}$ se define el conjunto de nivel de la función f como

$$N_c(f) = \{\boldsymbol{x} \in D : f(\boldsymbol{x}) = c\}$$

En el caso en que $n = 2$ y $n = 3$ el conjunto de nivel $N_c(f)$ se puede dibujar. Se le conoce como curva de nivel cuando $n = 2$ y superficie de nivel si $n = 3$.

Ejemplo 1.1.

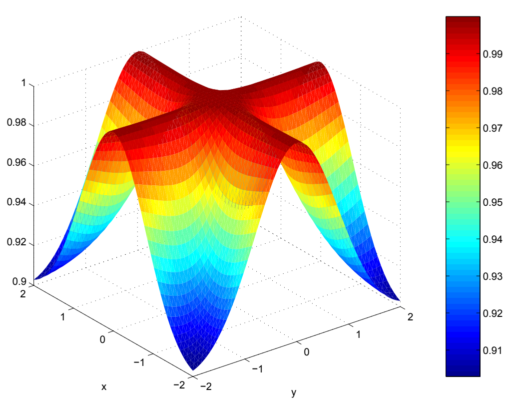


Figura 1.1: Grafo de sen $\left(\frac{xy}{x^2+y^2+1}\right)$

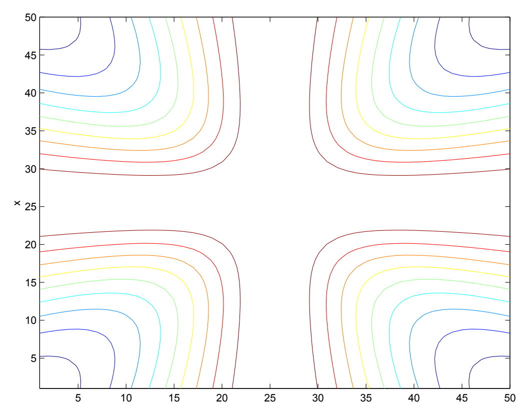


Figura 1.2: Curvas de nivel de sen $\left(\frac{xy}{x^2+y^2+1}\right)$

1.4. Conceptos introductorios de topología

La noción de límite y continuidad de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ involucra el carácter topológico de estos espacios, inducido por la norma.

En el estudio de la topología de $\mathbb{R}^n$, un rol fundamental es jugado por las bolas abiertas.

Definición 1.8. Dados $\boldsymbol{x}_0 \in \mathbb{R}^n$, $r \in \mathbb{R}_+$, llamaremos bola abierta de centro en $\boldsymbol{x}_0$ y radio r al conjunto

$$B(\boldsymbol{x}_0, r) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| < r\}$$

Llamaremos bola cerrada de centro en $\boldsymbol{x}_0$ y radio r al conjunto

$$\overline{B}(\boldsymbol{x}_0, r) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}$$

Definición 1.9. Diremos que $A \subseteq \mathbb{R}^n$ es un conjunto abierto si

$$(\forall \boldsymbol{x}_0 \in A)(\exists r > 0) : B(\boldsymbol{x}_0, r) \subseteq A$$

Ejemplo 1.2. $\mathbb{R}^n$ y el conjunto vacío $\varnothing$, son conjuntos abiertos. Aún cuando $\mathbb{R}^n$ es obviamente abierto, el caso del conjunto vacío requiere una reflexión. Si $\varnothing$ no es abierto entonces existe $\boldsymbol{x}_0 \in \varnothing$ tal que para todo $r > 0$, $B(\boldsymbol{x}_0, r) \cap \mathbb{R}^n \neq \varnothing$. Esto es absurdo pues no hay elementos en $\varnothing$.

Ejemplo 1.3. El conjunto $A = \{(x, y) : x > 1\}$ es un conjunto abierto. En efecto, si $(x, y) \in A$ entonces $B((x, y), \frac{x-1}{2}) \subset A$.

Ejemplo 1.4. Si $x_0 \in \mathbb{R}$ y $r > 0$ entonces $B(x_0, r)$ es un conjunto abierto. En efecto, si $x \in B(x_0, r)$ entonces $B(x, (r - \|x - x_0\|)/2) \subset B(x_0, r)$. Usando la desigualdad triangular muestre la veracidad de esta última afirmación y haga un dibujo.

Definición 1.10. Diremos que $A \subseteq \mathbb{R}^n$ es un conjunto cerrado si A^C es abierto.

Nota 1.2. Los conjuntos $\mathbb{R}^n$ y $\varnothing$ son abiertos y cerrados. También notamos que hay conjuntos que no son abiertos ni cerrados. Ver ejemplo a continuación.

Ejemplo 1.5. Sean $A = \{\frac{1}{n} : n \in \mathbb{N}\}$ y $B = A \cup \{0\}$. Entonces:

1. A no es cerrado. En efecto A^C no es abierto, pues $0 \in A^C$ y: $(\forall r > 0)$ $B(0, r) \not\subseteq A^C$. Lo anterior pues cualquiera sea $r > 0$ se tiene que

$$(\exists n \in \mathbb{N}) \text{ tal que } \frac{1}{n} < r$$

   Es decir, $\frac{1}{n} \in B(0, r)$.

2. A no es abierto pues $1 \in A$, pero $(\forall r > 0)$ $B(1, r) \not\subseteq A$.

3. B es cerrado y no es abierto.

Al igual que en el caso de $\mathbb{R}$, uno puede definir sucesiones de vectores en $\mathbb{R}^n$. Se trata de funciones de $\mathbb{N} \to \mathbb{R}^n$ tales que asignan un vector $\boldsymbol{x}_k$ a cada $k \in \mathbb{N}$. Usualmente se considera la notación $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}}$.

Definición 1.11. Una sucesión $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}}$ en $\mathbb{R}^n$ se dice sucesión convergente a $\boldsymbol{x} \in \mathbb{R}^n$ si:

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \ : \ \|\boldsymbol{x}_k - \boldsymbol{x}\| < \varepsilon$$

En el caso que $\boldsymbol{x} \in \mathbb{R}^n$ converge a $\boldsymbol{x}$ anotamos $\boldsymbol{x}_k \to \boldsymbol{x}$ o $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$. Notemos que la condición $\|\boldsymbol{x}_k - \boldsymbol{x}\| < \varepsilon$ es equivalente a $\boldsymbol{x}_k \in B(\boldsymbol{x}, \varepsilon)$.

Una proposición interesante, que se muestra a continuación, es que uno puede caracterizar los conjuntos cerrados mediante el uso de sucesiones. En realidad uno podría describir completamente la topología de $\mathbb{R}^n$ usando sucesiones, pero no lo haremos.

Proposición 1.4. $A \subseteq \mathbb{R}^n$ es cerrado si y sólo si:

$$\text{Para toda sucesión } \{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subseteq A : (\,(\boldsymbol{x}_k \to \boldsymbol{x}) \Rightarrow \boldsymbol{x} \in A)$$

Demostración.
($\Rightarrow$): Sea A un conjunto cerrado y sea $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subseteq A$ una sucesión cualquiera tal que $\lim\limits_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$. Queremos demostrar que $\boldsymbol{x} \in A$. Supongamos que esto no es cierto, es decir, que $\boldsymbol{x} \in A^C$. Como A es cerrado, existe $\varepsilon > 0$ tal que $B(\boldsymbol{x}, \varepsilon) \subset A^C$. Sin embargo, de la definición de convergencia, existe $k_0 \in \mathbb{N}$ tal que $\boldsymbol{x}_k \in B(\boldsymbol{x}, \varepsilon)$ para todo $k \geq k_0$, lo que implica que $x_k \in A^C$. Esto es una contradicción pues $\boldsymbol{x}_k \in A$ para todo $k \in \mathbb{N}$.

($\Leftarrow$): Para demostrar la recíproca probemos la contrarrecíproca. Es decir, si A no es cerrado entonces existe una sucesión $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subset A$ que converge a $\boldsymbol{x}$ y tal que $\boldsymbol{x} \notin A$.

Como A no es cerrado, A^C no es abierto, entonces existe un punto $\boldsymbol{x} \in A^C$ tal que

$$\text{Para todo } \varepsilon > 0 \text{ se tiene } B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing$$

Esta proposición nos permite construir una sucesión $\{\boldsymbol{x}_k\} \subset A$ de la siguiente manera: para cada $k \in \mathbb{N}$ tomamos $\varepsilon = \frac{1}{k}$ entonces, como $B(\boldsymbol{x}, 1/k) \cap A \neq \varnothing$, podemos elegir $\boldsymbol{x}_k \in B(\boldsymbol{x}, \frac{1}{k})$ y $\boldsymbol{x}_k \in A$. Por definición esta sucesión converge a $\boldsymbol{x}$, concluyendo la demostración pues $\boldsymbol{x} \notin A$. ■

Continuando con nuestra discusión sobre la topología de $\mathbb{R}^n$ hacemos algunas nuevas definiciones.

Definición 1.12. Sea $A \subseteq \mathbb{R}^n$. Entonces:

1. $\boldsymbol{x} \in A$ se dice punto interior de A si: $(\exists \varepsilon > 0),\ B(\boldsymbol{x}, \varepsilon) \subseteq A$.
2. $\boldsymbol{x} \in \mathbb{R}^n$ se dice punto adherente de A si: $(\forall \varepsilon > 0),\ B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing$.
3. $\boldsymbol{x} \in \mathbb{R}^n$ se dice punto de acumulación de A si: $(\forall \varepsilon > 0),\ (B(\boldsymbol{x}, \varepsilon) \setminus \{\boldsymbol{x}\}) \cap A \neq \varnothing$.
4. $\boldsymbol{x} \in \mathbb{R}^n$ se dice punto frontera de A si: $(\forall \varepsilon > 0),\ B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing \ \wedge\ B(\boldsymbol{x}, \varepsilon) \cap A^C \neq \varnothing$.

Observamos que un punto adherente de A no necesita estar en A, así mismo un punto de acumulación y un punto frontera no necesitan estar en A.

Definición 1.13. Dado $A \subseteq \mathbb{R}^n$ se definen los siguientes conjuntos:

1. Interior de A: $\mathrm{int}(A) = \{\boldsymbol{x} \in A : \boldsymbol{x} \text{ es punto interior de } A\}$.
2. Adherencia de A: $\mathrm{adh}(A) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \text{ es punto adherente de } A\}$.
3. Derivado de A: $\mathrm{der}(A) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \text{ es punto de acumulacion de } A\}$.
4. Frontera de A: $\partial A = \mathrm{fr}(A) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \text{ es punto frontera de } A\}$.

Nota 1.3. $\mathrm{int}(A) \subset A$ y $A \subset \mathrm{adh}(A)$.

Ejemplo 1.6. En $\mathbb{R}$ sea $A = [1, 2) \cup \{3\}$. Entonces: $\mathrm{der}(A) = [1, 2]$, $\mathrm{int}(A) = (1, 2)$, $\mathrm{adh}(A) = [1, 2] \cup \{3\}$ y $\mathrm{fr}(A) = \{1, 2, 3\}$.

Se obtiene de las definiciones la siguiente proposición:

Proposición 1.5. A es abierto si y sólo si $A = \mathrm{int}(A)$ y A es cerrado si y sólo si $A = \mathrm{adh}(A)$.

Ejemplo 1.7. Demuestre que $\boldsymbol{x} \in \text{adh}(A)$ si y sólo si existe una sucesión $\{\boldsymbol{x}_k\}_k \subset A$ tal que $\text{lím}_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$.

Solución.
($\Rightarrow$) Sea $\boldsymbol{x} \in \text{adh}(A)$, entonces se tiene que

$$B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing \ \forall \varepsilon > 0$$

Tomando ε de la forma $\varepsilon = 1/k$, se obtiene que:

$$B(\boldsymbol{x}, 1/k) \cap A \neq \varnothing \ \forall k \in \mathbb{N}$$

lo que implica que para todo $k \in \mathbb{N}$ existe un vector $\boldsymbol{x}_k$ perteneciente a $B(\boldsymbol{x}, 1/k) \cap A$. Entonces, $\{\boldsymbol{x}_k\}$ está contenida en A y $\text{lím}_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$.

($\Leftarrow$) Sea $x \in \mathbb{R}^n$ para el cual existe $\{\boldsymbol{x}_k\}_k \subset A$ tal que $\text{lím}_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$. Esto quiere decir que:

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \ : \ \|\boldsymbol{x}_k - \boldsymbol{x}\| < \varepsilon$$

lo que es equivalente a

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \ : \ \boldsymbol{x}_k \in B(\boldsymbol{x}, \varepsilon)$$

Pero $\boldsymbol{x}_k \in A$, entonces $\forall \varepsilon > 0 \ B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing$, lo cual es la definición de $\boldsymbol{x} \in \text{adh}(A)$. □

Con este ejemplo terminamos esta breve introdución a la topología de $\mathbb{R}^n$ y estamos en condiciones de presentar la noción de límite de una función.

1.5. Límites y continuidad

1.5.1. Límite de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Definición 1.14. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{a} \in \text{der}(A)$. Entonces decimos que $\boldsymbol{b}$ es el límite de f cuando $\boldsymbol{x}$ tiende a $\boldsymbol{a}$ y escribimos $\underset{\boldsymbol{x}\to\boldsymbol{a}}{\text{lím}} f(\boldsymbol{x}) = \boldsymbol{b}$ si

$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ tal que } (0 < \|\boldsymbol{x} - \boldsymbol{a}\| \leq \delta) \wedge (\boldsymbol{x} \in A) \Rightarrow \|f(\boldsymbol{x}) - \boldsymbol{b}\| \leq \varepsilon$$

Nota 1.4. En la anterior definición pedimos que $\boldsymbol{a} \in \text{der}(A)$ para que siempre haya puntos cerca de $\boldsymbol{a}$.

A continuación desarrollaremos dos ejemplos en los cuales debemos efectivamente demostrar el valor de un cierto límite. Para ello es necesario dar una fórmula que permita determinar δ dado ε.

Ejemplo 1.8. Demuestre usando la definición de límite que

$$\underset{(x,y)\to(1,1)}{\text{lím}} (x^2 - 1) = 0$$

Solución. En primer lugar veamos que si

$$\|(x, y) - (1, 1)\| \leq \delta$$

entonces necesariamente se cumple que

$$|x-1| \leq \delta \text{ e } |y-1| \leq \delta$$

por ende,

$$|x-1| \leq \delta \text{ y } |x+1| \leq \delta + 2 \tag{*}$$

Además, en este caso

$$|f(x)-b| = |x^2-1| = |x-1||x+1| \tag{**}$$

Así, para cualquier $\varepsilon > 0$ podemos elegir $\delta > 0$ que satisfaga

$$\delta(\delta+2) \leq \varepsilon$$

Entonces, si $\|(x,y)-(1,1)\| \leq \delta$, usando (*) y (**) se obtiene que

$$|f(x)-b| \leq \delta(\delta+2) \leq \varepsilon$$

$\square$

El ejemplo anterior es bastante simple. La dificultad puede aumentar cuando la función f es más complicada, por ejemplo si es un cuociente.

Ejemplo 1.9. Demuestre usando la definición que

$$\lim_{(x,y)\to(2,1)} \frac{x^2}{x-y} = 4$$

Solución. Partimos como en el ejemplo anterior viendo que si

$$\|(x,y)-(2,1)\| \leq \delta$$

entonces se tiene directamente que

$$|x-2| \leq \delta \text{ e } |y-1| \leq \delta \tag{*}$$

Por otro lado, un desarrollo algebraico simple nos lleva a lo siguiente

$$|f(x)-b| = \left|\frac{x^2}{x-y} - 4\right| = \frac{|x^2-4x+4y|}{|x-y|} = \frac{|(x-2)^2+4(y-1)|}{|x-y|} \tag{**}$$

Observamos aquí dos hechos relevantes. Primero, en el numerador tenemos una expresión que depende esencialmente de $|x-2|$ y $|y-1|$, cantidades controladas por δ, según (*). Segundo, en el denominador tenemos $|x-y|$, que es una cantidad que podría hacerse muy pequeña o incluso anularse si uno no es cuidadoso.

Para tratar este último término procedemos en una primera etapa encontrando un valor δ_1 medio, que si bien no nos permitirá acotar por ε nos permitirá controlar $|x-y|$. Esta cantidad no se anula en $(2,1)$, de aquí vemos que si elegimos $\delta_1 = 1/4$, por ejemplo, entonces trabajando un poco las desigualdades de (*) podemos concluir que

$$|x-y| > \frac{1}{2} \tag{***}$$

Supongamos ahora que se satisface (***), entonces de (*) y (**) obtenemos que

$$|f(x) - b| = \frac{|(x-2)^2 + 4(y-1)|}{|x-y|} \leq 2|(x-2)^2 + 4(y-1)|$$

De aquí, usando nuevamente (*) podemos acotar mejor

$$|f(x) - b| \leq 2|x-2|^2 + 8|y-1| \leq 2\delta^2 + 8\delta \leq 2\delta + 8\delta = 10\delta$$

Pues para $\delta < 1$ se tiene que $\delta^2 \leq \delta$. Entonces, dado $\varepsilon > 0$, para que se cumpla $|f(x) - b| \leq \varepsilon$ podemos elegir $\delta_2 = \frac{\varepsilon}{10}$. Sin embargo recordemos que lo anterior será cierto solamente si se satisface (***), para lo cual necesitamos que $\delta \leq \frac{1}{4}$. Por lo tanto, para asegurarnos que todos los argumentos anteriores sean válidos, elegimos $\delta = \min\{\frac{1}{4}, \frac{\varepsilon}{10}\}$ y se cumplirá que

$$\|(x,y) - (2,1)\| \leq \delta \Rightarrow \left| \frac{x^2}{x-y} - 4 \right| \leq \varepsilon$$

□

Ejemplo 1.10. Demostrar que

$$\lim_{(x,y)\to(0,0)} \frac{\operatorname{sen}(x^2+y^2)}{x^2+y^2} = 1$$

Solución. Del curso de Cálculo sabemos que

$$\lim_{\alpha\to 0} \frac{\operatorname{sen}(\alpha)}{\alpha} = 1$$

Entonces, dado $\varepsilon > 0$ existe $\delta_1 > 0$ tal que

$$|\alpha| \leq \delta_1 \Rightarrow \left| \frac{\operatorname{sen}(\alpha)}{\alpha} - 1 \right| \leq \varepsilon$$

Elijamos ahora $\delta = \sqrt{\delta_1}$. Entonces $\|(x,y) - (0,0)\| \leq \delta$ implica que $|x^2 + y^2| \leq \delta_1$ y entonces

$$\left| \frac{\operatorname{sen}(x^2+y^2)}{x^2+y^2} - 1 \right| \leq \varepsilon$$

□

Nota 1.5. Notamos que una manera equivalente de escribir la noción de límite es la siguiente: para toda bola abierta $B(\boldsymbol{b}, \varepsilon)$ existe una bola abierta $B(\boldsymbol{a}, \delta)$ tal que

$$\boldsymbol{x} \in (B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}) \cap A \Rightarrow f(\boldsymbol{x}) \in B(\boldsymbol{b}, \varepsilon)$$

o equivalentemente

$$f((B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}) \cap A) \subset B(\boldsymbol{b}, \varepsilon)$$

Ahora vamos a desarrollar el concepto de límite estudiando sus principales propiedades. En toda esta sección hacemos notar la analogía que se tiene con el curso de Cálculo, donde se estudio el concepto de límite y continuidad para funciones de una variable real. Es sorprendente que la mayoría de las proposiciones que veremos a continuación tienen una demostración que se obtiene de la análoga de Cálculo reemplazando $|\cdot|$ por $\|\cdot\|$.

Proposición 1.6. (Unicidad del límite)
Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{a} \in \mathrm{der}(A)$. Si

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}_1 \text{ y } \lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}_2$$

Entonces $\boldsymbol{b}_1 = \boldsymbol{b}_2$.

Demostración. Por hipótesis, dado $\varepsilon > 0$ existen $\delta_1 > 0$ y $\delta_2 > 0$ tales que

$$(\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \delta_1 \ \wedge \ \boldsymbol{x} \in A) \Rightarrow \|f(\boldsymbol{x}) - \boldsymbol{b}_1\| \leq \varepsilon$$

$$(\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \delta_2 \ \wedge \ \boldsymbol{x} \in A) \Rightarrow \|f(\boldsymbol{x}) - \boldsymbol{b}_2\| \leq \varepsilon$$

Entonces, si $\|\boldsymbol{x} - \boldsymbol{a}\| \leq \mathrm{mín}\{\delta_1, \delta_2\}$ se tendrá que

$$\|\boldsymbol{b}_1 - \boldsymbol{b}_2\| = \|(f(\boldsymbol{x}) - \boldsymbol{b}_1) - (f(\boldsymbol{x}) - \boldsymbol{b}_2)\| \leq \|f(\boldsymbol{x}) - \boldsymbol{b}_1\| + \|f(\boldsymbol{x}) - \boldsymbol{b}_2\| \leq 2\varepsilon$$

Como ε puede ser arbitrariamente pequeño, debemos tener que $\boldsymbol{b}_1 = \boldsymbol{b}_2$. ■

La siguiente proposición es muy útil para estudiar la existencia de un determinado límite. Se usa principalmente para mostrar que una cierta función no tiene límite.

Proposición 1.7. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{a} \in \mathrm{der}(A)$ tal que $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}$. Entonces para todo $B \subseteq A$ tal que $\boldsymbol{a} \in der(B)$ se tiene que

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

Demostración. Propuesto. ■

La contrarrecíproca nos da el siguiente corolario

Corolario 1.1. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{a} \in der(A)$ y $B_1, B_2 \subseteq A$, de manera que $\boldsymbol{a} \in der(B_1) \cap der(B_2)$. Si uno de los límites $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f|_{B_1}(\boldsymbol{x})$ ó $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f|_{B_2}(\boldsymbol{x})$ no existe o si $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f|_{B_1}(\boldsymbol{x}) \neq \lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f|_{B_2}(\boldsymbol{x})$ entonces el límite*

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

no existe.

Ejemplo 1.11. Se trata de estudiar el siguiente límite

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

Solución. Notemos en primer lugar que $f(0,0)$ no está definida, sin embargo esto no tiene importancia alguna. Lo que sí importa es que $(0,0) \in der(\mathbb{R}^2 \setminus \{(0,0)\})$ y que f está definida en $\mathbb{R}^2 \setminus \{(0,0)\}$. Si consideramos los conjuntos $A = \{(0,y) : y \neq 0\}$ y $B = \{(x,0) : x \neq 0\}$ entonces tenemos que

$$\lim_{(x,y)\to(0,0)} f|_A(x,y) = 0 \text{ y } \lim_{(x,y)\to(0,0)} f|_B(x,y) = 1$$

Concluimos entonces que el límite bajo estudio no existe, en virtud del corolario. □

Nota 1.6. En los ejemplos anteriores hemos considerado solamente funciones a valores reales, es decir, con espacio de llegada $\mathbb{R}$. Esto queda justificado en la siguiente proposición.

Proposición 1.8. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Entonces:

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b} \quad \Leftrightarrow \quad (\forall i = 1, \dots, m) \lim_{\boldsymbol{x}\to\boldsymbol{a}} f_i(\boldsymbol{x}) = b_i$$

Aquí f_i es la i-ésima función coordenada y b_i is la i-ésima coordenada de $\boldsymbol{b}$.

Demostración.
($\Rightarrow$): Propuesto.

($\Leftarrow$): Dado $\varepsilon > 0$, por hipótesis, para cualquier $i = 1, 2, ..., m$ existe $\delta_i > 0$ tal que $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \delta_i$ y $\boldsymbol{x} \in A$ implica que $|f_i(\boldsymbol{x}) - b_i| \leq \varepsilon/\sqrt{m}$. Entonces, si elegimos $\delta = \text{mín}\{\delta_1, ..., \delta_m\}$, tenemos que $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \delta$ y $\boldsymbol{x} \in A$ implica

$$\sum_{i=1}^{m}(f_i(\boldsymbol{x}) - b_i)^2 < m \cdot \frac{\varepsilon^2}{m} = \varepsilon^2$$

de donde $\|f(\boldsymbol{x}) - \boldsymbol{b}\| \leq \varepsilon$. ∎

Los siguientes tres teoremas resumen las propiedades importantes de los límites.

Teorema 1.11. *Sean $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{a} \in der(A)$. Si existen los límites $\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \boldsymbol{b}_1$ y $\lim_{\boldsymbol{x}\to\boldsymbol{a}} g(\boldsymbol{x}) = \boldsymbol{b}_2$, entonces los siguientes límites existen y pueden calcularse como sigue:*

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} (f + g)(\boldsymbol{x}) = \lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) + \lim_{\boldsymbol{x}\to\boldsymbol{a}} g(\boldsymbol{x}) = \boldsymbol{b_1} + \boldsymbol{b_2} \tag{1.1}$$

Si $\lambda \in \mathbb{R}$,

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} (\lambda f)(\boldsymbol{x}) = \lambda \lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = \lambda \boldsymbol{b_1} \tag{1.2}$$

Demostración. Demostraremos una a una las propiedades:

1. Dado $\varepsilon > 0$, como f y g tienen límite en $\boldsymbol{a}$, existen $\delta_1, \delta_2 > 0$ tales que

$$\begin{aligned} \boldsymbol{x} \in A, \|\boldsymbol{x} - \boldsymbol{a}\| \leq \delta_1 &\Rightarrow \|f(\boldsymbol{x}) - \boldsymbol{b_1}\| \leq \frac{\varepsilon}{2} \\ \boldsymbol{x} \in A, \|\boldsymbol{x} - \boldsymbol{a}\| \leq \delta_2 &\Rightarrow \|g(\boldsymbol{x}) - \boldsymbol{b_2}\| \leq \frac{\varepsilon}{2} \end{aligned}$$

   Definiendo $\delta = \text{mín}\{\delta_1, \delta_2\}$ se obtiene:

$$\boldsymbol{x} \in A, \ \|\boldsymbol{x} - \boldsymbol{a}\| \leq \delta \ \Rightarrow$$

$$\|(f + g)(\boldsymbol{x}) - (\boldsymbol{b_1} + \boldsymbol{b_2})\| \leq \|f(\boldsymbol{x}) - \boldsymbol{b_1}\| + \|g(\boldsymbol{x}) - \boldsymbol{b_2}\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

2. Si $\lambda = 0$ el resultado es trivial. Si $\lambda \neq 0$, dado $\varepsilon > 0$ como f tiene límite en $\boldsymbol{a}$, existe $\delta > 0$ tal que

$$\boldsymbol{x} \in A, \ \|\boldsymbol{x} - \boldsymbol{a}\| \leq \delta \ \Rightarrow \ \|f(\boldsymbol{x}) - \boldsymbol{b_1}\| \leq \frac{\varepsilon}{|\lambda|}$$

   lo que es equivalente a $\|\lambda f(\boldsymbol{x}) - \lambda \boldsymbol{b_1}\| \leq \varepsilon$ lo que concluye la demostración.

∎

Teorema 1.12. *Sean $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $g : f(A) \subseteq \mathbb{R}^m \to \mathbb{R}^p$. Si existen los límites* $\lim\limits_{\substack{\boldsymbol{x}\to\boldsymbol{a}\\ \boldsymbol{x}\in A}} f(\boldsymbol{x}) = \boldsymbol{b}$ *y* $\lim\limits_{\substack{\boldsymbol{y}\to\boldsymbol{b}\\ \boldsymbol{y}\in f(A)}} g(\boldsymbol{y}) = \boldsymbol{c}$*, entonces existe el límite*

$$\lim_{\boldsymbol{x}\to\boldsymbol{a},\ \boldsymbol{x}\in A} (g\circ f)(\boldsymbol{x}) = \boldsymbol{c} \tag{1.3}$$

Demostración. Dado $\varepsilon > 0$, como el límite de g en $\boldsymbol{b}$ existe, entonces existe un valor $\eta > 0$ tal que

$$(\|\boldsymbol{y}-\boldsymbol{b}\| \le \eta \ \wedge\ \boldsymbol{y}\in f(A)) \ \Rightarrow\ \|g(\boldsymbol{y})-\boldsymbol{c}\| \le \varepsilon$$

y como el límite de f en $\boldsymbol{a}$ existe, entonces existe un valor $\delta > 0$ tal que

$$(\|\boldsymbol{x}-\boldsymbol{a}\| \le \delta \ \wedge\ \boldsymbol{x}\in A) \ \Rightarrow\ \|f(\boldsymbol{x})-\boldsymbol{b}\| \le \eta$$

De estos dos resultados se concluye (usando $\boldsymbol{y} = f(\boldsymbol{x})$)

$$(\|\boldsymbol{x}-\boldsymbol{a}\| \le \delta \ \wedge\ \boldsymbol{x}\in A) \ \Rightarrow\ \|g(f(\boldsymbol{x}))-\boldsymbol{c}\| \le \varepsilon$$

por lo que $g\circ f$ tiene límite en $\boldsymbol{a}$. ∎

Teorema 1.13. *Sean $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Si existen los límites* $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = b_1$ *y* $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} g(\boldsymbol{x}) = b_2$*, entonces los siguientes límites existen y pueden calcularse como sigue:*

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} (f\cdot g)(\boldsymbol{x}) = \lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) \cdot \lim_{\boldsymbol{x}\to\boldsymbol{a}} g(\boldsymbol{x}) \tag{1.4}$$

Si $f(\boldsymbol{a}) \neq 0$

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \left(\frac{1}{f}\right)(\boldsymbol{x}) = \frac{1}{\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x})} \tag{1.5}$$

Demostración. Demostraremos una a una las propiedades y utilizaremos los dos teoremas anteriores.

1. Como $(f\cdot g)(\boldsymbol{x}) = \frac{1}{2}[(f+g)^2 - f^2 - g^2](\boldsymbol{x})$, para demostrar la existencia del límite de $f\cdot g$ en $\boldsymbol{a}$, la demostración se reduce al hecho de que el cuadrado de una función cuyo límite existe en $\boldsymbol{a}$ también tiene límite en $\boldsymbol{a}$.

   Debemos demostrar f^2 tiene límite en $\boldsymbol{a}$. Como f^2 es la composición de $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con $g : \mathbb{R} \to \mathbb{R}$ definida por $g(y) = y^2$, y como el límite de ambas funciones existe en $\boldsymbol{a}$ y $f(\boldsymbol{a})$ respectivamente, f^2 tiene límite en $\boldsymbol{a}$.

2. Siguiendo el desarrollo de la parte anterior, tenemos que $1/f$ es la composición de $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con $g : \mathbb{R}\setminus\{0\} \to \mathbb{R}$, definida por $g(y) = 1/y$, cuyos límites existen en $\boldsymbol{a}$ y $f(\boldsymbol{a})$ respectivamente.

∎

1.5.2. Continuidad de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Ahora que hemos estudiado el concepto de límite estamos preparados para introducir el concepto de continuidad de una función.

Definición 1.15. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y sea $\boldsymbol{x} \in A$. Decimos que f es continua en $\boldsymbol{x}$ si:

$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ tal que } [\boldsymbol{y} \in A \ \wedge \ \|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta] \ \Rightarrow \ \|f(\boldsymbol{x}) - f(\boldsymbol{y})\| \leq \varepsilon$$

o equivalentemente

$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ tal que } f(B(\boldsymbol{x}, \delta) \cap A) \subset B(f(\boldsymbol{x}), \varepsilon)$$

Definición 1.16. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y sea $\boldsymbol{x}_0 \in A$. Decimos que f es continua en A si:

$$(\forall \boldsymbol{x} \in A)(\forall \varepsilon > 0)(\exists \delta > 0) \text{ tal que } [\boldsymbol{y} \in A \ \wedge \ \|\boldsymbol{x} - \boldsymbol{y}\| \leq \delta] \ \Rightarrow \ \|f(\boldsymbol{x}) - f(\boldsymbol{y})\| \leq \varepsilon$$

o equivalentemente f es continua en todo punto de A.

Nota 1.7. Si $\boldsymbol{x}_0 \in A$ es un punto aislado de A, es decir, $\boldsymbol{x}_0 \in A \setminus \operatorname{der}(A)$ entonces f es obviamente continua en $\boldsymbol{x}_0$.

Si $\boldsymbol{x}_0 \in \operatorname{der}(A)$ entonces f es continua en $\boldsymbol{x}_0$ si y sólo si

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0)$$

Como consecuencia de la observación anterior y de la proposición 1.8 se tiene que $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es continua en $\boldsymbol{x}_0 \in A$ si y sólo si f_i es continua en $\boldsymbol{x}_0$, para todo $i = 1, \ldots, m$.

Como consecuencia de los teoremas 1.11, 1.12 y 1.13 tenemos también tres resultados sobre funciones continuas cuya demostración es análoga a lo que ya vimos sobre límites.

Teorema 1.14. *Sean* $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x}_0 \in A$ *y* $\lambda \in \mathbb{R}$. *Entonces:*

1. f *continua en* $\boldsymbol{x}_0$ *implica* λf *continua en* $\boldsymbol{x}_0$.
2. f *y* g *continuas en* $\boldsymbol{x}_0$ *implica* $f + g$ *continua en* $\boldsymbol{x}_0$.

Demostración. Propuesto. ■

Teorema 1.15. *Sean* $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $g : B \subseteq \mathbb{R}^m \to \mathbb{R}^p$, *con* $f(A) \subseteq B$ *y* $\boldsymbol{x}_0 \in A$. *Entonces* f *y* g *continuas en* $\boldsymbol{x}_0$ *implica* $g \circ f$ *continua en* $\boldsymbol{x}_0$.

Demostración. Sea $\varepsilon > 0$. Como g es continua en $f(\boldsymbol{x}_0)$ existe $\delta > 0$ tal que:

$$\boldsymbol{y} \in B, \ \|\boldsymbol{y} - f(\boldsymbol{x}_0)\| \leq \delta \ \Rightarrow \ \|g(\boldsymbol{y}) - g(f(\boldsymbol{x}_0))\| \leq \varepsilon$$

Por otro lado, de la continuidad de f en $\boldsymbol{x}_0$ sabemos que existe $\eta > 0$ tal que:

$$\boldsymbol{x} \in A, \ \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \eta \ \Rightarrow \ \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leq \delta$$

Juntando estas dos proposiciones y tomando $\boldsymbol{y} = f(\boldsymbol{x})$ se tiene que existe $\eta > 0$ tal que

$$\begin{aligned} &\boldsymbol{x} \in A, \ \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \eta \ \Rightarrow \ f(\boldsymbol{x}) \in B, \ \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leq \delta \\ &\Rightarrow \ \|g(f(\boldsymbol{x})) - g(f(\boldsymbol{x}_0))\| \leq \varepsilon \end{aligned}$$

Es decir:

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \eta \ \Rightarrow \ \|g \circ f(\boldsymbol{x}) - g \circ f(\boldsymbol{x}_0)\| \leq \varepsilon$$

■

Teorema 1.16. *Sean $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}$ y $\boldsymbol{x}_0 \in A$. Entonces f y g continuas en $\boldsymbol{x}_0$ implica $f \cdot g$ continua en $\boldsymbol{x}_0$. Agregando el supuesto $f(\boldsymbol{x}_0) \neq 0$, f continua en $\boldsymbol{x}_0$ implica $\frac{1}{f}$ continua en $\boldsymbol{x}_0$.*

Demostración. Propuesto. ■

Ejemplo 1.12. La función $f(x, y) = (\frac{x^2}{1+y^2}, x + y, y \cdot \operatorname{sen}(x))$ es continua en todo punto de $\mathbb{R}^2$.

Solución. En efecto, sus tres componentes son continuas en todo $\mathbb{R}^2$:

$f_1(x, y) = \frac{x^2}{1+y^2}$ lo es pues $f(x) = x^2$ lo es, al igual que $g(y) = 1 + y^2$ la cual además no se anula. Entonces, por el teorema anterior $\frac{1}{g(y)} = \frac{1}{1+y^2}$ también es continua, y nuevamente por el teorema anterior $x^2 \frac{1}{1+y^2}$ es continua.

$f_2(x, y) = x + y$ es evidentemente continua pues $f(x) = x$ e $g(y) = y$ son continuas y la suma de funciones continuas es continua.

$f_3(x) = \operatorname{sen}(x)$ es continua, $g(y) = y$ también lo es, luego por el teorema anterior $y \cdot \operatorname{sen}(x)$ es continua.

La idea es que con la ayuda de los teoremas 1.14, 1.15 y 1.16 podamos determinar por inspección cuando una función es continua, por ser una combinación de funciones continuas conocidas. □

Ejemplo 1.13. $f(x, y, z, t) = \operatorname{sen}(t(x^2 + y^2 + z^2))$ es continua en todo punto de $\mathbb{R}^4$.

Definición 1.17. Diremos que $A \subseteq \mathbb{R}^n$ es una vecindad de $\boldsymbol{x}_0 \in \mathbb{R}^n$ si $\boldsymbol{x}_0 \in A$ y A es abierto.

La siguiente proposición da una caracterización de la continuidad en términos de vecindades. Notar que la bola abierta $B(\boldsymbol{x}, r)$ es una vecindad de $\boldsymbol{x}$.

Proposición 1.9. $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es continua en $\boldsymbol{x}_0 \in \mathbb{R}^n$ si y sólo si para toda vecindad $U \subset \mathbb{R}^n$ de $f(\boldsymbol{x}_0)$ existe una vecindad $V \subset \mathbb{R}^n$ de $\boldsymbol{x}_0$ tal que $f(V \cap A) \subset U$.

Demostración.
($\Rightarrow$): Sea U una vecindad de $f(\boldsymbol{x}_0)$. Entonces existe $\varepsilon > 0$ tal que $B(f(\boldsymbol{x}_0), \varepsilon) \subset U$. Usando ahora la continuidad de f en $\boldsymbol{x}_0$, para este $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \delta \text{ y } \boldsymbol{x} \in A \Rightarrow \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leq \varepsilon$$

Es decir

$$\boldsymbol{x} \in B(\boldsymbol{x}_0, \delta) \text{ y } \boldsymbol{x} \in A \Rightarrow \; f(\boldsymbol{x}) \in B(f(\boldsymbol{x}_0), \varepsilon)$$

Si definimos $V = B(\boldsymbol{x}_0, \delta)$, que es una vecindad de $\boldsymbol{x}_0$, obtenemos de aquí que

$$f(B(\boldsymbol{x}_0, \delta) \cap A) = f(V \cap A) \subset U$$

($\Leftarrow$): Esta implicancia queda propuesta. ■

1.6. Diferenciabilidad de funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$

Consideremos $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ y $\boldsymbol{x} = (x_1, \ldots, x_j, \ldots, x_n) \in A$, donde el conjunto A se supone abierto a fin de evitar dificultades con elección de la dirección sobre la cual definir el concepto de derivada parcial. Para $1 \leq j \leq n$ fijo, definimos la función:

$$\begin{aligned} \phi : \mathbb{R} &\to \mathbb{R} \\ h &\mapsto f(\boldsymbol{x} + h\boldsymbol{e}_j) \end{aligned}$$

donde $\boldsymbol{e}_j$ es el elemento j-ésimo de la base canónica de $\mathbb{R}^n$. Notamos que $\boldsymbol{x}+h\boldsymbol{e}_j = (x_1, \ldots, x_{j-1}, h+x_j, x_{j+1}, \ldots, x_n)$. A esta función, siendo de $\mathbb{R}$ en $\mathbb{R}$, le podemos aplicar la teoría de diferenciabilidad desarrollada en el curso de Cálculo.

Definición 1.18. Llamaremos derivada parcial de f con respecto a x_j en $\boldsymbol{x} \in \mathbb{R}^n$ a

$$\frac{\partial f}{\partial x_j}(\boldsymbol{x}) = \lim_{h\to 0} \frac{\phi(h)-\phi(0)}{h} = \lim_{h\to 0} \frac{f(\boldsymbol{x}+h\boldsymbol{e}_j)-f(\boldsymbol{x})}{h},$$

cuando dicho límite existe. Si la derivada parcial de f con respecto a x_j existe en todo punto de A, entonces ella define una función

$$\frac{\partial f}{\partial x_j} : A \subseteq \mathbb{R}^n \to \mathbb{R}$$

Nota 1.8. Como una derivada parcial es una derivada de una función de $\mathbb{R}$ en $\mathbb{R}$, uno puede usar todas las reglas de derivación estudiadas en Cálculo.

Ejemplo 1.14. Calcular la derivada parcial de la función f con respecto a x para

$$f(x,y) = x^4 y + \operatorname{sen}(xy)$$

Solución. Haciendo uso de la nota anterior tenemos que

$$\frac{\partial f}{\partial x}(x) = 4x^3 y + \cos(xy)y$$

□

Ejemplo 1.15. Calcular la derivada parcial de la función f con respecto a x para $f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}$.

Solución. Para esto notamos que si $(x,y) \neq (0,0)$ entonces

$$\frac{\partial f}{\partial x}(x) = \frac{y^3}{(x^2+y^2)^{\frac{3}{2}}}$$

Si definimos $f(0,0) = 0$ entonces podemos calcular también la derivada parcial de f con respecto a x en $(0,0)$. Aquí usamos la definición

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{0}{h} = 0$$

□

En la noción de diferenciabilidad hay dos aspectos a considerar: el aspecto analítico y el aspecto geométrico. Parece atractivo y simple definir una función diferenciable en un punto como una función que posee todas sus derivadas parciales en dicho punto. Sin embargo esta condición no es suficiente para producir una buena definición, pues nos gustaría que al menos se cumplan las siguientes propiedades:

1. Criterio geométrico. Si f es diferenciable en un punto entonces es pueda definir un plano tangente al grafo de la función en dicho punto.

2. Criterio analítico. La composición de funciones diferenciables debe ser diferenciable.

Ejemplo 1.16. El siguiente ejemplo ilustra la importancia de los requisitos anteriores. Consideremos la función definida por $f(x,y) = x^{\frac{1}{3}}y^{\frac{1}{3}}$. Calculando las derivadas parciales por definición obtenemos que estas existen y son

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$$

Uno esperaría que el plano tangente a la función f en el punto $(0,0,0)$ estuviese dado por la ecuación $z = 0$, para ser consecuentes con el valor de las derivadas parciales de f en el $(0,0)$. Sin embargo, este plano no puede ser tangente al grafo de la función en $(0,0)$, pues sobre la recta $y = x$ el grafo de f tiene pendiente infinita en el origen.

Por otro lado, $g : \mathbb{R} \to \mathbb{R}^2$ definida por $g(x) = (x,x)$ es diferenciable en cero con $g_1'(0) = g_2'(0) = 1$, pero $f \circ g(x) = x^{\frac{2}{3}}$ no es diferenciable en 0.

Concluimos de este ejemplo que la noción de diferenciabilidad debe involucrar algo más que la sola existencia de las derivadas parciales en el punto.

Veamos ahora el aspecto geométrico en $\mathbb{R}^2$ para motivar la definición en general. Supongamos que queremos ajustar un plano tangente al grafo de $f : \mathbb{R}^2 \to \mathbb{R}$ en el punto $(x_0, y_0, f(x_0,y_0)) \in G(f)$. Es decir queremos encontrar un plano en $\mathbb{R}^3$ que pase por el punto y que tenga la misma inclinación que el grafo de f en el punto.

Para esto consideremos un plano genérico $z = a + bx + cy$ y deduzcamos de las condiciones que mencionamos antes el valor de las constantes a, b y c. En primer lugar queremos que (x_0, y_0, z_0) pertenezca al plano tangente, entonces

$$f(x_0, y_0) = z_0 = a + bx_0 + cy_0$$

Fijando la variable x_0 se debería tener que las recta $z = a + bx_0 + cy$ debe ser tangente al grafo de $z = f(x_0, y)$ en el punto y_0. Esto implica que

$$\frac{\partial f}{\partial x}(x_0, y_0) = b$$

De manera análoga

$$\frac{\partial f}{\partial y}(x_0, y_0) = c$$

Es decir, el plano buscado es

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

Definición 1.19. Sea $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$, A abierto y $(x_0, y_0) \in A$. Decimos que f es diferenciable en (x_0, y_0) si existen las derivadas parciales $\frac{\partial f}{\partial x}(x_0, y_0)$ y $\frac{\partial f}{\partial y}(x_0, y_0)$ y el límite

$$\lim_{(x,y)\to(x_0,y_0)} \frac{|f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0)|}{\|(x,y) - (x_0,y_0)\|} = 0$$

existe.

Intuitivamente estamos pidiendo a f, para que sea diferenciable, que su grafo y el plano tangente al grafo en el punto estén muy cerca, tan cerca que incluso al dividir su distancia por $\|(x,y)-(x_0,y_0)\|$ la fracción todavía es pequeña.

Vamos ahora al caso general:

Definición 1.20. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, A abierto y $\boldsymbol{x}_0 \in A$. Entonces, si todas las derivadas parciales de todas las funciones coordenadas existen en $\boldsymbol{x}_0$, podemos definir una matriz $Df(\boldsymbol{x}_0) \in M_{m\times n}(\mathbb{R})$ como

$$(Df(\boldsymbol{x}_0))_{ij} = \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)$$

Esta matriz se conoce como matriz Jacobiana o Diferencial de f en $\boldsymbol{x}_0$.

Definición 1.21. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, A abierto y $\boldsymbol{x}_0 \in A$. Decimos que f es diferenciable en $\boldsymbol{x}_0$ si para todo $i = 1, \ldots, m$ y para todo $j = 1, \ldots, n$ $\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)$ existe y

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0 \tag{1.6}$$

Usualmente la condición (1.6) se escribe de manera equivalente como

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{\|f(\boldsymbol{x}_0+\boldsymbol{h}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)\boldsymbol{h}\|}{\|\boldsymbol{h}\|} = 0 \tag{1.7}$$

Nota 1.9. En vista de la proposición 1.8 tenemos que una función $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $\boldsymbol{x}_0 \in A$ si y sólo si cada una de las funciones componentes $f_i : A \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en $\boldsymbol{x}_0 \in A$.

La diferenciabilidad tiene que ver con la posibilidad de aproximar la función $f(\boldsymbol{x})$ por una función lineal afín $L(\boldsymbol{x}) = f(\boldsymbol{x}_0) + Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)$ cerca de $\boldsymbol{x}_0$. En ocasiones L se llama aproximación de primer orden de f en $\boldsymbol{x}_0$. Como ya vimos en el caso de $n = 2$ y $m = 1$, lo anterior se interpreta geométricamente como la posibilidad de ajustar el plano tangente al grafo de f en $\boldsymbol{x}_0$.

Ejemplo 1.17. Encuentre el plano tangente al grafo de $f(x,y) = x^2 + y^2$ en el punto $(1,1)$.

Solución. Calculamos las derivadas parciales en $(1,1)$

$$\frac{\partial f}{\partial x}(1,1) = 2x|_{(1,1)} = 2 \qquad y \qquad \frac{\partial f}{\partial y}(1,1) = 2y|_{(1,1)} = 2$$

Entonces el plano tangente tiene por ecuación

$$z = 2 + 2(x-1) + 2(y-1)$$

es decir

$$z = -2 + 2x + 2y$$

□

1.6.1. Aproximación de primer orden

Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es una función diferenciable en $\boldsymbol{x}_0 \in A$. Entonces la función lineal afín

$$L(\boldsymbol{h}) = f(\boldsymbol{x}_0) + Df(\boldsymbol{x}_0)\boldsymbol{h}$$

es una aproximación de la función $f(\boldsymbol{x}_0 + \boldsymbol{h})$ cerca de $\boldsymbol{h} = 0$. No sólo $f(\boldsymbol{x}_0 + \boldsymbol{h})$ y $L(\boldsymbol{h})$ son muy parecidas, sino que además

$$\lim_{\substack{\boldsymbol{h}\to\boldsymbol{0}\\ \boldsymbol{h}\neq\boldsymbol{0}}} \frac{f(\boldsymbol{x}_0 + \boldsymbol{h}) - L(\boldsymbol{h})}{\|\boldsymbol{h}\|} = 0 \tag{1.8}$$

En el caso $m = 1$ y $n = 2$ tenemos que

$$L(\boldsymbol{h}) = L(h_1, h_2) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h_1 + \frac{\partial f}{\partial y}(x_0, y_0)h_2$$

aproxima a f al primer orden cerca de (x_0, y_0).

Ejemplo 1.18. Encuentre una aproximación de primer orden para la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = x\cos(y) + e^{xy}$, cerca del punto $(x_0, y_0) = (1, \pi)$.

Solución. Tenemos $f(1, \pi) = -1 + e^\pi$ y

$$\begin{aligned} \frac{\partial f}{\partial x}(1, \pi) &= -1 + \pi e^\pi \\ \frac{\partial f}{\partial y}(1, \pi) &= e^\pi \end{aligned}$$

Así la aproximación de primer orden es

$$f(1 + a, \pi + b) \approx -1 + e^\pi + (-1 + \pi e^\pi)a + e^\pi b$$

□

Nota 1.10. En general, los vectores de $\mathbb{R}^n$ deberían considerarse como vectores columna. De esta manera la matriz Jacobiana y el producto $Df(\boldsymbol{x}_0)\boldsymbol{h}$, $\boldsymbol{h} \in \mathbb{R}^n$ tienen sentido. Sin embargo, cuando esto no lleve a confusión, muchas veces escribiremos los vectores de $\mathbb{R}^n$ como vectores fila, por economía notacional.

1.6.2. Gradiente de una función

Definición 1.22. La matriz Jacobiana de una función de $\mathbb{R}^n$ en $\mathbb{R}$ es una matriz de $1 \times n$

$$Df(\boldsymbol{x}_0) = \left[\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\ \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0) \dots \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0)\right]$$

Es costumbre identificar esta matriz con un vector de $\mathbb{R}^n$ llamado gradiente de f en $\boldsymbol{x}_0$. Así

$$\nabla f(\boldsymbol{x}_0) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \dots, \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0)\right)^T$$

Con esta notación la aproximación lineal tiene la siguiente forma

$$L(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)$$

1.6.3. Relación entre continuidad y diferenciabilidad

A continuación abordamos la relación entre continuidad y diferenciabilidad. Al igual que en el caso de las funciones de una variable, la diferenciabilidad es un concepto más fuerte que el de continuidad, en el sentido que, toda función diferenciable en un punto es también continua en dicho punto. Pero la relación no termina allí. Si bien la mera existencia de las derivadas parciales de una función en un punto no garantiza su diferenciabilidad, en el caso que esas derivadas parciales existen y son continuas en una vecindad de dicho punto entonces sí hay diferenciabilidad de la función.

Comencemos extendiendo la desigualdad de Cauchy-Schwarz al producto de una matriz por un vector. Sea A una matriz en $M_{m\times n}(\mathbb{R})$ y $\boldsymbol{x} \in \mathbb{R}^n$. Entonces, usando la desigualdad de Cauchy-Schwarz obtenemos

$$\|A\boldsymbol{x}\|^2 = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 \leq \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}^2 \sum_{j=1}^{n} x_j^2\right) = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)\|\boldsymbol{x}\|^2 \tag{1.9}$$

Definición 1.23. La desigualdad 1.9 sugiere definir la norma de una matriz $A \in M_{m\times n}(\mathbb{R})$ como

$$\|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} \tag{1.10}$$

(1.10) se conoce como norma Frobenius.

Podemos resumir esto en la siguiente proposición

Proposición 1.10. Si $A \in M_{m\times n}(\mathbb{R})$ y $\boldsymbol{x} \in \mathbb{R}^n$ entonces

$$\|A\boldsymbol{x}\| \leq \|A\|\,\|\boldsymbol{x}\|$$

Teorema 1.17. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{x}_0 \in A$. Entonces*

$$f \text{ es diferenciable en } \boldsymbol{x}_0 \;\Rightarrow\; f \text{ es continua en } \boldsymbol{x}_0$$

Demostración. Si f es diferenciable en $\boldsymbol{x}_0$ entonces existe la matriz Jacobiana de f en $\boldsymbol{x}_0$ y

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} = 0$$

Así dado $\varepsilon > 0$, existe $\delta_1 > 0$ tal que si $\|\boldsymbol{x}-\boldsymbol{x}_0\| < \delta_1$ entonces:

$$\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\| < \varepsilon\|\boldsymbol{x}-\boldsymbol{x}_0\|$$

Entonces, usando la desigualdad triangular y la proposición 1.10, obtenemos

$$\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leq \varepsilon\|\boldsymbol{x}-\boldsymbol{x}_0\| + \|Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\| \leq (\varepsilon + \|Df(\boldsymbol{x}_0)\|)\|\boldsymbol{x}-\boldsymbol{x}_0\|$$

De esta manera, si elegimos $\delta = min\{\delta_1, \varepsilon/(\varepsilon + \|Df(\boldsymbol{x}_0)\|)\}$ obtenemos

$$\|\boldsymbol{x}-\boldsymbol{x}_0\| < \delta \Rightarrow \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| < \varepsilon$$

es decir, f es continua en $\boldsymbol{x}_0$. ■

Corolario 1.2. *Si f es diferenciable en $\boldsymbol{x}_0$ entonces existen δ y K tales que*

$$\|\boldsymbol{x}-\boldsymbol{x}_0\| < \delta \quad\Rightarrow\quad \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| \leq K\|\boldsymbol{x}-\boldsymbol{x}_0\|$$

Demostración. Basta tomar $\varepsilon = 1$ y $K = 1 + \|Df(\boldsymbol{x}_0)\|$ en la demostración del teorema. ■

Retomaremos esta propiedad cuando veamos más adelante el teorema del valor medio.

Continuando con la relación entre diferenciabilidad y continuidad consideremos una función $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $\boldsymbol{x}_0 \in A$. Supongamos que las derivadas parciales de f existen no sólo en $\boldsymbol{x}_0$ sino que en una vecindad de $V \subseteq A$ de $\boldsymbol{x}_0$. Entonces estas definen funciones

$$\frac{\partial f_i}{\partial x_j} : V \to \mathbb{R} \text{ tal que } \boldsymbol{x} \mapsto \frac{\partial f_i}{\partial x_j}(\boldsymbol{x})$$

El siguiente teorema nos da una condición suficiente para que una función f sea diferenciable en $\boldsymbol{x}_0$.

Teorema 1.18. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ tal que sus derivadas parciales existen en una vecindad de $\boldsymbol{x}_0$ y son continuas en $\boldsymbol{x}_0$. Entonces f es diferenciable en $\boldsymbol{x}_0$.*

Para la demostración vamos a utilizar el teorema 1.7.

Demostración. En vista de la nota 1.9, basta demostrar el teorema para $m = 1$. La demostración en el caso general para n es análoga al caso $n = 2$. Haremos la demostración para el caso $n = 2$ y el caso general quedará propuesto.

Consideremos

$$\begin{gathered} f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) = \\ f(x_{01} + h_1, x_{02} + h_2) - f(x_{01} + h_1, x_{02}) + f(x_{01} + h_1, x_{02}) - f(x_{01}, x_{02}) \end{gathered}$$

Fijemos $x_{01} + h_1$ y supongamos que $h_2 > 0$. Definamos la función $g : [0, h_2] \to \mathbb{R}$ como $g(s) = f(x_{01} + h_1, x_{02} + s)$. Como la derivada parcial con respecto a y existe en una vecindad de $\boldsymbol{x}_0$, si $\boldsymbol{h}$ es pequeño, la función g es derivable en $[0, h_2]$. Entonces, aplicando el teorema del valor medio (teorema 1.7) a g existe $c_2 \in [x_{02}, x_{02} + h_2]$ tal que

$$f(x_{01} + h_1, x_{02} + h_2) - f(x_{01} + h_1, x_{02}) = \frac{\partial f}{\partial x_2}(x_{01} + h_1, x_{02} + c_2)$$

Si $h_2 < 0$ entonces se define g en $[h_2, 0]$ y se procede de manera análoga. Notemos que en cualquier caso $|c_2| < |h_2| \leq \|\boldsymbol{h}\|$.

Por el mismo argumento, fijando ahora x_{02} se tendrá que existe c_1 tal que $|c_1| < |h_1| \leq \|\boldsymbol{h}\|$ y

$$f(x_{01} + h_1, x_{02}) - f(x_{01}, x_{02}) = \frac{\partial f}{\partial x_1}(x_{01} + c_1, x_{02})$$

Y así, si anotamos $\boldsymbol{x}_1^* = (x_{01} + c_1, x_{02})$ y $\boldsymbol{x}_2^* = (x_{01} + h_1, x_{02} + c_2)$ tenemos

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_1^*)h_1 + \frac{\partial f}{\partial x_2}(\boldsymbol{x}_2^*)h_2$$

y de aquí

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)\boldsymbol{h} = \left[\frac{\partial f}{\partial x_1}(\boldsymbol{x}_1^*) - \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\right] h_1 + \left[\frac{\partial f}{\partial x_2}(\boldsymbol{x}_2^*) - \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\right] h_2 \qquad (*)$$

Usemos ahora la continuidad de las derivadas parciales en $\boldsymbol{x}_0$. Dado $\varepsilon > 0$ existe $\delta_1 > 0$ tal que

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta_1 \Rightarrow \left| \frac{\partial f}{\partial x_1}(\boldsymbol{x}) - \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0) \right| < \frac{\varepsilon}{\sqrt{2}}$$

y existe $\delta_2 > 0$ tal que

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta_2 \Rightarrow \left| \frac{\partial f}{\partial x_2}(\boldsymbol{x}) - \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0) \right| < \frac{\varepsilon}{\sqrt{2}}$$

Eligiendo $\delta = min\{\delta_1, \delta_2\}$, si $\|\boldsymbol{h}\| < \delta$ entonces para $k = 1, 2$,

$$\|\boldsymbol{x}_k^* - \boldsymbol{x}_0\| = |c_k| < |h_k| \leq \|\boldsymbol{h}\| < \delta$$

y así

$$\left| \frac{\partial f}{\partial x_j}(\boldsymbol{x}_k^*) - \frac{\partial f}{\partial x_j}(\boldsymbol{x}_0) \right| < \frac{\varepsilon}{\sqrt{2}}$$

Usando la desigualdad de Cauchy-Schwarz en la ecuación (*), se obtiene

$$|f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)\boldsymbol{h}| \leq \sqrt{\left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_1^*) - \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\right)^2 + \left(\frac{\partial f}{\partial x_2}(\boldsymbol{x}_2^*) - \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\right)^2} \cdot \|\boldsymbol{h}\|$$

de donde

$$|f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)\boldsymbol{h}| \leq \sqrt{\frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2}} \cdot \|\boldsymbol{h}\| = \varepsilon\|\boldsymbol{h}\|$$

■

Definición 1.24. f se dice de clase $\mathcal{C}^1$ en $\boldsymbol{x}_0 \in A$ si f es diferenciable en $\boldsymbol{x}_0$ y todas las derivadas parciales de f son continuas en $\boldsymbol{x}_0$.

Definición 1.25. $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ se dice diferenciable en A si f es diferenciable en cada punto de A. f se dice de clase $\mathcal{C}^1$ en A si f es diferenciable en A y todas sus derivadas parciales son continuas en A.

Ejemplo 1.19. $f(x, y) = \frac{cos(x)+e^{xy}}{x^2+y^2}$ es diferenciable en cada punto $(x, y) \neq (0, 0)$ pues $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ existen y son continuas en cualquier punto distinto del $(0, 0)$.

Ejemplo 1.20. Las derivadas parciales de la función $f(x, y) = x^{1/3}y^{1/3}$ estudiada anteriormente no estan definidas en una vecindad del origen, menos pueden ser continuas.

Volvemos ahora a la idea de aproximación que lleva el concepto de diferenciabilidad. Vimos que si f es diferenciable en un punto $\boldsymbol{x}_0$ entonces la función lineal afín $L(\boldsymbol{h}) = f(\boldsymbol{x}_0) + Df(\boldsymbol{x}_0)\boldsymbol{h}$ aproxima a f en primer orden. Nos interesa ahora la recíproca.

Teorema 1.19. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función continua en $\boldsymbol{x}_0 \in A$ y L una función lineal afín, $L(\boldsymbol{h}) = \boldsymbol{a} + B\boldsymbol{h}$ con $\boldsymbol{a} \in \mathbb{R}^m$ y $B \in M_{m\times n}(\mathbb{R})$. Supongamos que*

$$\lim_{\boldsymbol{h}\to 0} \frac{\|f(\boldsymbol{x}_0 + \boldsymbol{h}) - L(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0 \tag{1.11}$$

Entonces $\boldsymbol{a} = f(\boldsymbol{x}_0)$ y $B = Df(\boldsymbol{x}_0)$ y f es diferenciable en $\boldsymbol{x}_0$.

Demostración. Por hipótesis

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{\|f(\boldsymbol{x}_0+\boldsymbol{h})-L(\boldsymbol{h})\|}{\|\boldsymbol{h}\|}=0$$

entonces se tiene en particular que

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} f(\boldsymbol{x}_0+\boldsymbol{h})-\boldsymbol{a}-B\boldsymbol{h}=0$$

Como f es continua en $\boldsymbol{x}_0$, esto implica que $f(\boldsymbol{x}_0)=\boldsymbol{a}$.

Por otra parte, para cada $1\leq j\leq n$, tenemos de la hipótesis que se cumple lo siguiente (en este límite h es un número real y $h\boldsymbol{e}_j$ un vector que tiende a cero)

$$\lim_{h\to 0} \frac{\|f(\boldsymbol{x}_0+h\boldsymbol{e}_j)-f(\boldsymbol{x}_0)-Bh\boldsymbol{e}_j\|}{\|h\boldsymbol{e}_j\|}=0$$

es decir,

$$\lim_{h\to 0} \frac{f(\boldsymbol{x}_0+h\boldsymbol{e}_j)-f(\boldsymbol{x}_0)}{h}=B\boldsymbol{e}_j$$

Así, las derivadas parciales de las funciones coordenadas con respecto a x_j existen y son iguales a la columna j-ésima de B. De aquí $B=Df(\boldsymbol{x}_0)$. Reemplazando estos valores en la hipótesis, tenemos que f es diferenciable en $\boldsymbol{x}_0$. ■

Este teorema dice que cuando existe una aproximación lineal de primer orden esta es única. Es un teorema de unicidad. Por otra parte este teorema es muy útil para demostrar la diferenciabilidad de una cierta función. La idea es que uno tiene una matriz B que es candidata a ser la matriz Jacobiana de f. Con dicha matriz uno demuestra (1.11). y concluye la diferenciabilidad.

El siguiente es un teorema que permite combinar funciones diferenciables.

Teorema 1.20. *Sean $f,g:A\subseteq\mathbb{R}^n\to\mathbb{R}^m$, $\boldsymbol{x}_0\in A$ funciones tales que $Df(\boldsymbol{x}_0)$ y $Dg(\boldsymbol{x}_0)$ existen, entonces*

1. *$D(f+g)(\boldsymbol{x}_0)$ existe y se tiene*

$$D(f+g)(\boldsymbol{x}_0)=Df(\boldsymbol{x}_0)+Dg(\boldsymbol{x}_0) \tag{1.12}$$

2. *Si $\lambda\in\mathbb{R}$, $D(\lambda f)(\boldsymbol{x}_0)$ existe y se tiene*

$$D(\lambda f)(\boldsymbol{x}_0)=\lambda Df(\boldsymbol{x}_0) \tag{1.13}$$

3. *Si $m=1$, $D(f\times g)(\boldsymbol{x}_0)$ existe y se tiene*

$$D(f\cdot g)(\boldsymbol{x}_0)=g(\boldsymbol{x}_0)Df(\boldsymbol{x}_0)+f(\boldsymbol{x}_0)Dg(\boldsymbol{x}_0) \tag{1.14}$$

4. *Si $m=1$ y $f(\boldsymbol{x}_0)\neq 0$, $D(1/f)(\boldsymbol{x}_0)$ existe y se tiene*

$$D(1/f)(\boldsymbol{x}_0)=-\frac{1}{f^2(\boldsymbol{x}_0)}Df(\boldsymbol{x}_0) \tag{1.15}$$

Demostración. Veamos una a una las propiedades:

1. (1.12) y (1.13) son consecuencia directa de la definición de diferenciabilidad.

2. Una forma de demostrar (1.14) es simplemente ver que se cumpla la ecuación (1.11). Es decir,

$$
\begin{aligned}
&\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{(f\cdot g)(\boldsymbol{x}_0+\boldsymbol{h})-(f\cdot g)(\boldsymbol{x}_0)-[g(\boldsymbol{x}_0)Df(\boldsymbol{x}_0)+f(\boldsymbol{x}_0)Dg(\boldsymbol{x}_0)]\boldsymbol{h}}{\|\boldsymbol{h}\|}\\
&=\lim_{\boldsymbol{h}\to\boldsymbol{0}}\left(\frac{f(\boldsymbol{x}_0+\boldsymbol{h})[g(\boldsymbol{x}_0+\boldsymbol{h})-g(\boldsymbol{x}_0)-Dg(\boldsymbol{x}_0)\boldsymbol{h}]}{\|\boldsymbol{h}\|}\right.\\
&\quad\left.+\frac{g(\boldsymbol{x}_0)[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h}]}{\|\boldsymbol{h}\|}+\frac{[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)]Dg(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}\right)\\
&=\lim_{\boldsymbol{h}\to\boldsymbol{0}} f(\boldsymbol{x}_0+\boldsymbol{h})\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{g(\boldsymbol{x}_0+\boldsymbol{h})-g(\boldsymbol{x}_0)-Dg(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}\\
&\quad+g(\boldsymbol{x}_0)\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}+\lim_{\boldsymbol{h}\to\boldsymbol{0}}[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)]\frac{Dg(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}=0
\end{aligned}
$$

Como ya vimos para la aproximación de primer orden, tenemos que el último límite se anula ya que $Dg(\boldsymbol{x}_0)\boldsymbol{h}/\|\boldsymbol{h}\|$ es acotado y $[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)]$ tiende a cero.

3. Para demostrar 1.15, nuevamente debemos verificar que se cumple (1.11).

$$
\begin{aligned}
&\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\frac{1}{f(\boldsymbol{x}_0+\boldsymbol{h})}-\frac{1}{f(\boldsymbol{x}_0)}+\frac{Df(\boldsymbol{x}_0)\boldsymbol{h}}{[f(\boldsymbol{x}_0)]^2}}{\|\boldsymbol{h}\|}=\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{f(\boldsymbol{x}_0)^2-f(\boldsymbol{x}_0+\boldsymbol{h})f(\boldsymbol{x}_0)+f(\boldsymbol{x}_0+\boldsymbol{h})Df(\boldsymbol{x}_0)\boldsymbol{h}}{f(\boldsymbol{x}_0+\boldsymbol{h})f(\boldsymbol{x}_0)^2\|\boldsymbol{h}\|}\\
&=\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{1}{f(\boldsymbol{x}_0+\boldsymbol{h})f(\boldsymbol{x}_0)^2}\cdot\frac{Df(\boldsymbol{x}_0)\boldsymbol{h}[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)]-f(\boldsymbol{x}_0)[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h}]}{\|\boldsymbol{h}\|}\\
&=\frac{1}{f(\boldsymbol{x}_0)^3}\lim_{\boldsymbol{h}\to\boldsymbol{0}}[f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)]\frac{Df(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}-\frac{1}{f(\boldsymbol{x}_0)^2}\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h}}{\|\boldsymbol{h}\|}=0
\end{aligned}
$$

■

Teorema 1.21. (Regla de la cadena)
Sean $f: A\subseteq\mathbb{R}^n\to\mathbb{R}^m$ y $g: B\subseteq\mathbb{R}^m\to\mathbb{R}^p$ tales que f es diferenciable en $\boldsymbol{x}_0\in A$ y g diferenciable en $f(\boldsymbol{x}_0)\in B$. Entonces $g\circ f: A\subset\mathbb{R}^n\to\mathbb{R}^p$ es diferenciable en $\boldsymbol{x}_0$ y

$$D(f\circ g)(\boldsymbol{x}_0)=Dg(f(\boldsymbol{x}_0))Df(\boldsymbol{x}_0)$$

Demostración. De acuerdo al teorema 1.19 basta demostrar que

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\|g(f(\boldsymbol{x}_0+\boldsymbol{h}))-g(f(\boldsymbol{x}_0))-Dg(f(\boldsymbol{x}_0))Df(\boldsymbol{x}_0)\boldsymbol{h}\|}{\|\boldsymbol{h}\|}=0$$

De la desigualdad triangular y de Cauchy-Schwarz tenemos que

$$
\begin{aligned}
&\|g(f(\boldsymbol{x}_0+\boldsymbol{h}))-g(f(\boldsymbol{x}_0))-Dg(f(\boldsymbol{x}_0))Df(\boldsymbol{x}_0)\boldsymbol{h}\|\\
\le&\|g(f(\boldsymbol{x}_0+\boldsymbol{h}))-g(f(\boldsymbol{x}_0))-Dg(f(\boldsymbol{x}_0))(f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0))\|\\
+&\|Dg(f(\boldsymbol{x}_0))\|\,\|(f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h})\|.
\end{aligned}
$$

Por otro lado, como f es diferenciable en $\boldsymbol{x}_0$, dado $\varepsilon>0$ existe $\delta_1>0$ tal que

$$\|\boldsymbol{h}\|<\delta_1\Rightarrow\|f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)-Df(\boldsymbol{x}_0)\boldsymbol{h}\|\le\frac{\varepsilon}{2\|Dg(f(\boldsymbol{x}_0))\|}\|\boldsymbol{h}\|$$

y de esta forma el segundo término de la suma queda acotado por $\frac{\varepsilon}{2}\|\boldsymbol{h}\|$. Del resultado anterior, usando la desigualdad triangular, también se deduce que existe $K>0$ tal que

$$\|\boldsymbol{h}\|<\delta_1\Rightarrow\|f(\boldsymbol{x}_0+\boldsymbol{h})-f(\boldsymbol{x}_0)\|\le K\|\boldsymbol{h}\|$$

Usando la diferenciabilidad de g en $f(\boldsymbol{x}_0)$ encontramos que existe $\delta_2 > 0$ tal que:

$$\|\boldsymbol{a}\| < \delta_2 \Rightarrow \|g(f(\boldsymbol{x}_0) + \boldsymbol{a}) - g(f(\boldsymbol{x}_0)) - Dg(f(\boldsymbol{x}_0))\boldsymbol{a}\| \leq \frac{\varepsilon}{2K}\|\boldsymbol{a}\|$$

Luego, escogiendo $\delta = min\{\delta_1, \frac{\delta_2}{K}\}$, si $\|\boldsymbol{h}\| < \delta$ y considerando $\boldsymbol{a} = f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0)$ tendremos

$$\begin{aligned}\|g(f(\boldsymbol{x}_0 + \boldsymbol{h})) - g(f(\boldsymbol{x}_0)) - Dg(f(\boldsymbol{x}_0))(f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0))\| &\leq \frac{\varepsilon}{2K}\|f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0)\| \\ &\leq \frac{\varepsilon}{2K}K\|\boldsymbol{h}\| \\ &= \frac{\varepsilon}{2}\|\boldsymbol{h}\|\end{aligned}$$

Y por lo tanto

$$\|\boldsymbol{h}\| < \delta \Rightarrow \frac{\|g(f(\boldsymbol{x}_0 + \boldsymbol{h})) - g(f(\boldsymbol{x}_0)) - Dg(f(\boldsymbol{x}_0))Df(\boldsymbol{x}_0)\boldsymbol{h}\|}{\|\boldsymbol{h}\|} \leq \varepsilon$$

La demostración requiere que $\|Dg(f(\boldsymbol{x}_0))\| > 0$. Indicar cómo hacerlo cuando es cero queda propuesto. ■

Ejemplo 1.21. (Caso particular)
Sean $g : \mathbb{R}^3 \to \mathbb{R}$ y $f : \mathbb{R}^3 \to \mathbb{R}^3$ tal que

$$f(x, y, z) = (u(x, y, z), v(x, y, z), w(x, y, z))$$

Sea $F(x, y, z) = g \circ f(x, y, z) = g(u(x, y, z), v(x, y, z), w(x, y, z))$. Calculemos $\frac{\partial F}{\partial x}(x, y, z)$.

Solución. Para esto usemos la regla de la cadena para obtener $DF(x, y, z)$

$$DF(x, y, z) = Dg(f(x, y, z))Df(x, y, z) = \left(\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}, \frac{\partial g}{\partial w}\right)\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{bmatrix}$$

Entonces

$$\frac{\partial F}{\partial x} = \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial x}$$

□

Ejemplo 1.22. (Coordenadas esféricas)
Sea $g : \mathbb{R}^3 \to \mathbb{R}$ y consideremos el cambio de variables a coordenadas esféricas

$$\begin{aligned} x &= r\cos(\theta)\operatorname{sen}(\phi) \\ y &= r\operatorname{sen}(\theta)\operatorname{sen}(\phi) \\ z &= r\cos(\phi) \end{aligned}$$

Definamos $F(r, \phi, \theta) = g(r\cos(\theta)\operatorname{sen}(\phi), r\operatorname{sen}(\theta)\operatorname{sen}(\phi), r\cos(\theta))$. Queremos calcular las derivadas parciales de F con respecto a r, θ y ϕ.
Si miramos el cambio de variables como una función $f : \mathbb{R}^3 \to \mathbb{R}^3$, usando el ejemplo anterior se

tendrá:

$$\begin{aligned}
\frac{\partial F}{\partial r} &= \frac{\partial g}{\partial x}\cos(\theta)\operatorname{sen}(\phi) + \frac{\partial g}{\partial y}\operatorname{sen}(\theta)\operatorname{sen}(\phi) + \frac{\partial g}{\partial z}\cos(\phi)\\
\frac{\partial F}{\partial \phi} &= \frac{\partial g}{\partial x} r\cos(\theta)\cos(\phi) + \frac{\partial g}{\partial y} r\operatorname{sen}(\theta)\cos(\phi) - \frac{\partial g}{\partial z} r\operatorname{sen}(\phi)\\
\frac{\partial F}{\partial \theta} &= -\frac{\partial g}{\partial x} r\operatorname{sen}(\theta)\operatorname{sen}(\phi) + \frac{\partial g}{\partial y} r\cos(\theta)\operatorname{sen}(\phi)
\end{aligned}$$

Ejemplo 1.23. Sea $h(x,y) = f(u(x,y), v(x,y))$ donde la función f esta definida por $f(u,v) = \frac{u^2+v^2}{u^2-v^2}$, y las funciones u y v por $u(x,y) = e^{-x-y}$ y $v(x,y) = e^{xy}$. Calcular $\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}$.

Solución. Usando la regla de la cadena tenemos

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = \frac{4uv^2}{(u^2-v^2)^2}e^{-x-y} + \frac{4vu^2}{(u^2-v^2)^2}ye^{xy}$$

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = \frac{4uv^2}{(u^2-v^2)^2}e^{-x-y} + \frac{4vu^2}{(u^2-v^2)^2}xe^{xy}$$

□

Ejemplo 1.24. (Importante)
Sean $f : \mathbb{R}^3 \to \mathbb{R}$, $\gamma : \mathbb{R} \to \mathbb{R}^3$ y $h : \mathbb{R} \to \mathbb{R}$ de manera que: $h(t) = f \circ \gamma(t)$. Entonces:

$$\frac{dh}{dt} = \frac{\partial f}{\partial x}\dot{\gamma}_1 + \frac{\partial f}{\partial y}\dot{\gamma}_2 + \frac{\partial f}{\partial z}\dot{\gamma}_3 = \nabla f(\gamma(t)) \cdot \dot{\gamma}(t)$$

La función γ representa una curva en el espacio y $\dot{\gamma}$ su vector tangente.

Ejemplo 1.25. Consideremos las funciones

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

$$g(u,v) = (u+v, u, v^2)$$

Calcular $D(g \circ f)(1,1)$.

Solución. Usemos la regla de la cadena y calculemos

$$Dg(f(1,1))Df(1,1)$$

$$Df(x,y) = \begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix} \text{ y evaluando } Df(1,1) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Evaluamos $f(1,1) = (2,1)$ y luego calculamos

$$Dg(u,v) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2v \end{bmatrix}$$

que al evaluar nos da

$$Dg(f(1,1)) = Dg(2,1) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Entonces tenemos

$$D(g \circ f)(1,1) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 4 \end{bmatrix}$$

Una forma alternativa, que nos da evidentemente el mismo resultado, es desarrollar h primero

$$h(x,y) = g \circ f(x,y) = (x^2 + y^2 + 1, x^2 + 1, y^4)$$

y luego derivar y evaluar

$$Dh(x,y) = \begin{bmatrix} 2x & 2y \\ 2x & 0 \\ 0 & 4y^3 \end{bmatrix} \text{ entonces } Dh(1,1) = \begin{bmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 4 \end{bmatrix}$$

□

Ejemplo 1.26. (Derivación implícita)
Sean $G : \mathbb{R}^2 \to \mathbb{R}$ e $y : \mathbb{R} \to \mathbb{R}$ tales que

$$G(x, y(x)) = 0 \qquad \forall x \in \mathbb{R} \text{ entonces}$$

$$\frac{\partial G}{\partial x} + \frac{\partial G}{\partial y}\frac{dy}{dx} = 0$$

Si $\frac{\partial G}{\partial y} \neq 0$ entonces podemos despejar la derivada de y con respecto a x

$$\frac{dy}{dx} = -\frac{\frac{\partial G}{\partial x}}{\frac{\partial G}{\partial y}}$$

Nota 1.11. Más adelante veremos que dada la ecuación

$$G(x,y) = 0$$

hay condiciones que garantizan la posibilidad de despejar y como función de x. Este criterio lo da el teorema de la función implícita, que veremos más adelante, y consiste en suponer que $\frac{\partial G}{\partial y}$ es no nula. Notamos que justamente este término aparece en el denominador de la derivada de y.

Ejemplo 1.27. Veamos ahora un caso de derivación implícita con más variables. Sean

$$G_1(x, y_1(x), y_2(x)) = 0$$

$$G_2(x, y_1(x), y_2(x)) = 0$$

donde $G_1, G_2 : \mathbb{R}^3 \to \mathbb{R}$ y $y_1, y_2 : \mathbb{R} \to \mathbb{R}$. Suponiendo que todas las funciones involucradas son diferenciables se desea calcular $\dot{y}_1(x)$ e $\dot{y}_2(x)$.

$$\frac{\partial G_1}{\partial x} + \frac{\partial G_1}{\partial y_1}\dot{y}_1 + \frac{\partial G_1}{\partial y_2}\dot{y}_2 = 0$$

$$\frac{\partial G_2}{\partial x} + \frac{\partial G_2}{\partial y_1}\dot{y}_1 + \frac{\partial G_2}{\partial y_2}\dot{y}_2 = 0$$

Este es un sistema de 2×2 que tiene la siguiente forma:

$$\begin{pmatrix} \frac{\partial G_1}{\partial y_1} & \frac{\partial G_1}{\partial y_2} \\ \frac{\partial G_2}{\partial y_1} & \frac{\partial G_2}{\partial y_2} \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = -\begin{pmatrix} \frac{\partial G_1}{\partial x} \\ \frac{\partial G_2}{\partial x} \end{pmatrix}$$

del cual podemos despejar las derivadas buscadas, bajo la hipótesis de invertibilidad correspondiente:

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = \frac{-1}{\frac{\partial G_1}{\partial y_1}\frac{\partial G_2}{\partial y_2} - \frac{\partial G_2}{\partial y_1}\frac{\partial G_1}{\partial y_2}} \begin{pmatrix} \frac{\partial G_2}{\partial y_2} & \frac{-\partial G_1}{\partial y_2} \\ \frac{-\partial G_2}{\partial y_1} & \frac{\partial G_1}{\partial y_1} \end{pmatrix} \begin{pmatrix} \frac{\partial G_1}{\partial x} \\ \frac{\partial G_2}{\partial x} \end{pmatrix}$$

Teorema 1.22. *Una condición necesaria y suficiente para que*

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y}$$

es que exista g tal que $f(x,y) = g(x+y)$.

Demostración. La suficiencia de dicha condición es evidente. Sólo debemos verificar que la condición también es necesaria. Para esto consideremos el siguiente cambio de variables:

$$u = x+y,\ v = x-y \quad \Leftrightarrow \quad x = \frac{u+v}{2},\ y = \frac{u-v}{2}$$

y definamos

$$h(u,v) = f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) = f(x,y)$$

Tendremos que

$$\frac{\partial h}{\partial v}(u,v) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = \frac{1}{2}\frac{\partial f}{\partial x} - \frac{1}{2}\frac{\partial f}{\partial y} = 0$$

Es decir h no depende de v, y por lo tanto:

$$f(x,y) = h(u) = h(x+y)$$

■

1.6.4. Teorema del valor medio en $\mathbb{R}^n$

El teorema del valor medio para funciones de un intervalo $[a,b]$ en $\mathbb{R}$ puede extenderse a varias variables. Esto haremos a continuación.

Teorema 1.23. (Teorema del valor medio en $\mathbb{R}^n$)
Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en A y $[\boldsymbol{a},\boldsymbol{b}] \subseteq A$, donde $[\boldsymbol{a},\boldsymbol{b}] = \{\boldsymbol{a} + t(\boldsymbol{b}-\boldsymbol{a}) : t \in [0,1]\}$. Entonces existe $\boldsymbol{c} \in [\boldsymbol{a},\boldsymbol{b}]$ tal que

$$\|f(\boldsymbol{b}) - f(\boldsymbol{a})\| = \nabla f(\boldsymbol{c}) \cdot (\boldsymbol{b}-\boldsymbol{a})$$

y por lo tanto

$$\|f(\boldsymbol{b}) - f(\boldsymbol{a})\| \le \sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]} \|Df(\boldsymbol{x})\|\|\boldsymbol{b}-\boldsymbol{a}\|$$

Demostración. Definamos la función $g : [0,1] \to \mathbb{R}$ por $g(t) = f(\boldsymbol{a}+t(\boldsymbol{b}-\boldsymbol{a}))$. Nótese que $g(0) = f(\boldsymbol{a})$ y $g(1) = f(\boldsymbol{b})$. Por definición g es derivable y por regla de la cadena se obtiene

$$g'(t) = \nabla f(\boldsymbol{a} + t(\boldsymbol{b}-\boldsymbol{a})) \cdot (\boldsymbol{b}-\boldsymbol{a})$$

Como la función g satisface las hipótesis del teorema 1.7 para funciones reales podemos concluir que existe $t_0 \in (0,1)$ tal que $g(1) - g(0) = g'(t_0)(1-0)$ y por lo tanto,

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = \nabla f(\boldsymbol{a} + t_0(\boldsymbol{b}-\boldsymbol{a})) \cdot (\boldsymbol{b}-\boldsymbol{a}) \tag{*}$$

Tomando módulo y usando la desigualdad de Cauchy-Schwarz

$$|f(\boldsymbol{b}) - f(\boldsymbol{a})| \leq \|\nabla f(\boldsymbol{a} + t_0(\boldsymbol{b}-\boldsymbol{a}))\| \, \|\boldsymbol{b}-\boldsymbol{a}\| \leq \left(\sup_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]} \|\nabla f(\boldsymbol{x})\| \right) \|\boldsymbol{b}-\boldsymbol{a}\|$$

■

Nota 1.12. El teorema anterior también se conoce como teorema de los incrementos finitos. Para una función con valores en $\mathbb{R}^m$, no necesariamente se obtiene una igualdad como en el caso de una variable o como en (*). Esto no es problema, pues su utilidad realmente viene del hecho que los incrementos se pueden acotar.

Si agregamos la hipótesis de que f es de clase $\mathcal{C}^1$ entonces $Df(\cdot)$ es continua y como $\|\cdot\|$ también lo es, $\|Df(\cdot)\|$ será composición de funciones continuas y por lo tanto continua de $\mathbb{R}$ en $\mathbb{R}$ y así alcanzará su máximo sobre un intervalo cerrado. Luego

$$\|f(\boldsymbol{b}) - f(\boldsymbol{a})\| \leq \max_{\boldsymbol{x}\in[\boldsymbol{a},\boldsymbol{b}]} \|Df(\boldsymbol{x})\| \|\boldsymbol{b}-\boldsymbol{a}\|$$

1.7. Gradiente y geometría

Recordemos que el gradiente $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ se define como

$$\nabla f(\boldsymbol{x}_0) = \left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0), \ldots, \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0) \right)^T$$

El gradiente se define generalmente sólo en puntos donde f es diferenciable, aún cuando basta que existan las derivadas parciales para determinarlo.

Ejemplo 1.28. Calcular el gradiente de $f(x,y,z) = \sqrt{x^2+y^2+z^2} = r$.

Solución.

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{2x}{2\sqrt{x^2+y^2+z^2}} = \frac{x}{r} \\
\frac{\partial f}{\partial y} &= \frac{2y}{2\sqrt{x^2+y^2+z^2}} = \frac{y}{r} \\
\frac{\partial f}{\partial z} &= \frac{2z}{2\sqrt{x^2+y^2+z^2}} = \frac{z}{r} \\
\nabla f(x,y,z) &= \frac{1}{\sqrt{x^2+y^2+z^2}}(x,y,z) = \widehat{r}
\end{aligned}$$

□

Consideremos a continuación un vector unitario $\boldsymbol{v}$ (con $\|\boldsymbol{v}\| = 1$) y miremos la función:

$$t \to f(\boldsymbol{x}_0 + t\boldsymbol{v})$$

Esta es una función de una variable y uno puede preguntarse sobre su diferenciabilidad en $t = 0$, es decir sobre la existencia del límite

$$\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t}$$

esto da lugar a la siguiente definición

Definición 1.26. Si el límite

$$\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t}$$

existe decimos que f tiene derivada direccional en $\boldsymbol{x}_0$, en la dirección $\boldsymbol{v}$. A este límite se lo anota usualmente como $Df(\boldsymbol{x}_0, \boldsymbol{v})$ o $\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$. Un caso particular de derivada direccional son las derivadas parciales.

Proposición 1.11. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es una función diferenciable, A abierto y $\boldsymbol{x}_0 \in A$ entonces f tiene derivadas direccionales bien definidas en cualquier dirección unitaria $\boldsymbol{v}$ y

$$Df(\boldsymbol{x}_0, \boldsymbol{v}) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} v_i$$

Demostración. Definamos $g : \mathbb{R} \to \mathbb{R}$ por

$$g(t) = f(\boldsymbol{x}_0 + t\boldsymbol{v})$$

entonces, por la definición de derivada tenemos que

$$g'(0) = \lim_{t\to 0} \frac{g(t) - g(0)}{t} = \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t} \Rightarrow g'(0) = Df(\boldsymbol{x}_0, \boldsymbol{v}) \tag{*}$$

Por otra parte, $g(t) = f(\boldsymbol{x})$ con $\boldsymbol{x} = (x_{01} + tv_1, \ldots, x_{0n} + tv_n)$ y aplicando la regla de la cadena se obtiene

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x})}{\partial x_i} \frac{\partial x_i}{\partial t} = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x})}{\partial x_i} v_i$$

para $t = 0$ se obtiene

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} \frac{\partial x_i}{\partial t} = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} v_i \tag{**}$$

De (*) y (**) se concluye

$$Df(\boldsymbol{x}_0, \boldsymbol{v}) = \sum_{i=1}^{n} \frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} v_i$$

y utilizando la notación del gradiente nos queda

$$Df(\boldsymbol{x}_0, \boldsymbol{v}) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v}$$

∎

La siguiente proposición nos sirve para fijar ideas y resume lo que ya se ha expuesto.

Proposición 1.12. Sea $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}$, con A abierto y $(0, 0) \in A$, una función continua y con derivadas parciales discontinuas en $(0, 0)$. Podemos dar las siguientes propiedades, de fácil verificación, que no son conclusivas sobre la diferenciabilidad de la función f en $(0, 0)$:

1. Derivadas parciales continuas en $(0,0) \Rightarrow$ diferenciabilidad en $(0,0)$.

   No nos sirve para concluir diferenciabilidad, pues las derivadas parciales no son continuas. Tampoco nos sirve para mostrar que no es diferenciable, pues se trata de una condición suficiente para diferenciabilidad, no de una condición necesaria.

2. Diferenciabilidad en $(0,0) \Rightarrow$ las derivadas parciales existen en $(0,0)$.

   No nos sirve para probar que la función no es diferenciable, pues las derivadas parciales existen en todos los puntos. Tampoco sirve para probar diferenciabilidad, ya que es una propiedad que caracteriza una condición necesaria y no suficiente para diferenciabilidad.

Una propiedad conclusiva sobre la diferenciabilidad es la siguiente: Si f es diferenciable en $(0,0)$ entonces f tiene derivadas direccionales bien definidas en $(0,0)$. Si se tiene alguna dirección, por ejemplo $\boldsymbol{v} = (1,0)$, en la que la derivada direccional no está bien definida entonces la función no es diferenciable en $(0,0)$.

Ejemplo 1.29. Consideremos la función

$$f(x,y) = x^{\frac{1}{2}} y^{\frac{1}{2}}$$

La función es continua en $(0,0)$ ya que si $(x_n, y_n) \to (0,0)$ se tiene

$$|f(x_n,y_n) - f(0,0)| = |x^{\frac{1}{2}}||y^{\frac{1}{2}}| \to 0$$

entonces

$$\lim_{n\to\infty} f(x_n,y_n) = f(0,0)$$

como la sucesión es arbitraria,

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

y se concluye que f es continua en $(0,0)$.

Para la continuidad de las derivadas parciales consideremos lo siguiente: Sea $\boldsymbol{v} = (v_1, v_2)$ tal que $\|\boldsymbol{v}\| = 1$, entonces

$$Df[(0,0);\boldsymbol{v}] = \lim_{t\to 0} \frac{f[(0,0)+t\boldsymbol{v}] - f(0,0)}{t} = v_1^{\frac{1}{2}} v_2^{\frac{1}{2}}$$

En particular, evaluando en $\boldsymbol{v} = (1,0)$ y $\boldsymbol{v} = (0,1)$ se obtienen respectivamente

$$\frac{\partial f}{\partial x}(0,0) = 0 \text{ y } \frac{\partial f}{\partial y}(0,0) = 0$$

De acuerdo al teorema 1.11, si f es diferenciable en $(0,0)$ se esperaría que para toda dirección unitaria $\boldsymbol{v}$

$$Df[(0,0);\boldsymbol{v}] = \nabla f(0,0) \cdot \boldsymbol{v} = 0$$

sin embargo, la derivada direccional obtenida nos dice que en la dirección $\boldsymbol{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ (nótese que $\|\boldsymbol{v}\| = 1$) se obtiene $Df[(0,0);\boldsymbol{v}] = \frac{1}{\sqrt{2}}$ lo que indica que f no es diferenciable en $(0,0)$.

El significado geométrico de la derivada direccional lo podemos ver en el siguiente ejemplo:

Ejemplo 1.30. Encontrar la derivada direccional $D_{\boldsymbol{v}}f(x,y)$ de $f(x,y) = x^3 - 3xy + 4y^2$ si $\boldsymbol{v}$ es el vector unitario dado por el ángulo $\alpha = \pi/4$. ¿Cuál es el valor de $D_{\boldsymbol{v}}f(1,2)$?

Solución. A partir de $D_{\boldsymbol{v}}f(x,y) = f_x \cos(\alpha) + f_y \operatorname{sen}(\alpha)$ se obtiene

$$D_{\boldsymbol{v}}f(x,y) = (3x^2 - 3y)\frac{1}{\sqrt{2}} + (-3x + 8y)\frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2}}[3(x^2 - x) + 5y]$$

Luego evaluamos en $(x_0, y_0) = (1,2)$

$$D_{\boldsymbol{v}}f(1,2) = 5\sqrt{2}$$

La derivada direccional $D_{\boldsymbol{v}}f(1,2)$ representa la razón de cambio en la dirección de $\boldsymbol{v}$. Ésta es la pendiente de la de la recta tangente a la curva de intersección de la superficie $z = x^3 - 3xy + 4y^2$ y el plano vertical que pasa por $(1,2,0)$ en la dirección de $\boldsymbol{v}$ como se ilustra en el siguiente gráfico

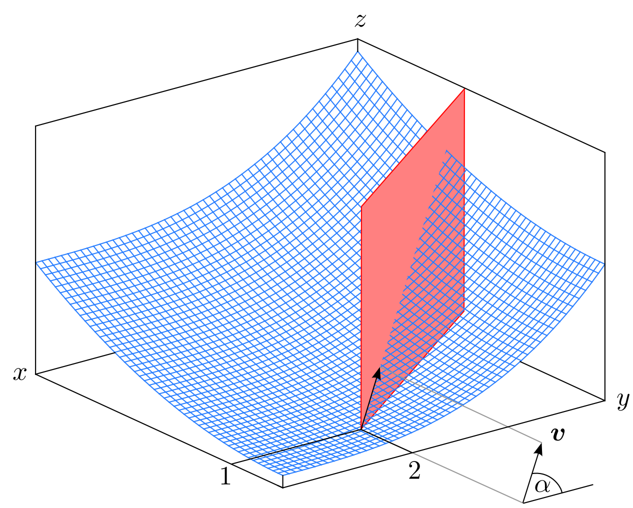


Figura 1.3: Interpretación geométrica de la derivada direccional

□

$Df(\boldsymbol{x}_0, \boldsymbol{v})$ corresponde a la derivada de la restricción de f a la recta

$$L : \boldsymbol{x} = \boldsymbol{x}_0 + t\boldsymbol{v}$$

En el caso en que f es diferenciable en $\boldsymbol{x}_0$ entonces la derivada direccional existe en cualquier dirección y se puede calcular como:

$$Df(\boldsymbol{x}_0, \boldsymbol{v}) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v}$$

Notemos que hemos considerado $\boldsymbol{v}$ unitario. Esto se hace para no distorsionar la escala espacial entre $f(\boldsymbol{x}_0)$ y $f(\boldsymbol{x}_0 + t\boldsymbol{v})$.

Ejemplo 1.31. Calcular la razón de cambio de la función f en la dirección $\boldsymbol{v} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ en el punto $(1,0,0)$ para

$$f(x,y,z) = x^2 e^{-yz}$$

Solución. Como f es diferenciable, se pide calcular $r = \nabla f(1,0,0) \cdot (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$. Para ello calculamos las derivadas parciales

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xe^{-yz} \\ \frac{\partial f}{\partial y} &= -x^2 z e^{-yz} \\ \frac{\partial f}{\partial z} &= -x^2 y e^{-yz} \end{aligned}$$

Evaluando obtenemos

$$\nabla f(1,0,0) = (2,0,0)^T \Rightarrow r = (2,0,0) \cdot (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) = \frac{2}{\sqrt{3}}$$

□

El siguiente teorema nos muestra un importante aspecto geométrico del gradiente.

Teorema 1.24. *Si $\nabla f(\boldsymbol{x}_0) \neq 0$ entonces $\nabla f(\boldsymbol{x}_0)$ apunta en la dirección en la cual f crece más rapidamente.*

Demostración. Sea $\boldsymbol{v}$ vector unitario, entonces la razón de cambio de f en la dirección de $\boldsymbol{v}$ es

$$\begin{aligned} \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v} &= \|\nabla f(\boldsymbol{x}_0)\| \|\boldsymbol{v}\| \cos\theta \\ &= \|\nabla f(\boldsymbol{x}_0)\| \cos\theta \end{aligned}$$

Donde θ es el ángulo entre $\boldsymbol{v}$ y $\nabla f(\boldsymbol{x}_0)$. Este ángulo es máximo cuando $\boldsymbol{v}$ y $\nabla f(\boldsymbol{x}_0)$ son paralelos. ■

Otro aspecto geométrico importante del gradiente viene en el estudio de superficies. Sea $F : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable y consideremos el conjunto

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : F(\boldsymbol{x}) = k\}$$

Este conjunto es, en general, una superficie si $n = 3$ y una hipersuperficie si $n > 3$.

El vector $\nabla F(\boldsymbol{x}_0)$ es ortogonal a la superficie en $\boldsymbol{x}_0$.

La idea geometrica es la siguiente: Supongamos que $c : \mathbb{R} \to \mathbb{R}^n$ es una función diferenciable, es decir es una curva suave en $\mathbb{R}^n$. Supongamos además que c está sobre S es decir:

$$F(c(t)) = k \qquad \forall t$$

Si $c(t_0) = \boldsymbol{x}_0$ entonces derivando y evaluando en $\boldsymbol{x}_0$

$$\nabla F(\boldsymbol{x}_0) \cdot c'(t_0) = 0$$

$c'(t_0)$ es una dirección tangente a la superficie, y como c es una curva cualquiera $\nabla F(\boldsymbol{x}_0)$ debe ser ortogonal al plano tangente a la superficie.

Definición 1.27. Sea $F : A \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable. Si $F(\boldsymbol{x}_0) = 0$, $\nabla F(\boldsymbol{x}_0) \neq 0$, se define el plano tangente a $S = \{\boldsymbol{x} \in A : F(\boldsymbol{x}) = 0\}$ en $\boldsymbol{x}_0$ como:

$$\nabla F(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0$$

Nota 1.13. Ciertamente, si $F(c(t)) = 0$ y $c(t_0) = \boldsymbol{x}_0$ entonces $c'(t_0) + \boldsymbol{x}_0$ está en el plano tangente.

Se puede demostrar que si una dirección $\boldsymbol{v}$ está en el plano tangente, entonces existe c tal que $\boldsymbol{x} = c'(t_0) + \boldsymbol{x}_0$ Esta propiedad geométrica es muy importante y su demostración la postergaremos hasta que hayamos demostrado el teorema de la función implícita.

Ejemplo 1.32. Encontrar el plano tangente a la superficie

$$x^2 + y^2 + z^2 - 1 = 0 \text{ en } (1, 0, 0)$$

Solución.

$$\nabla F(1, 0, 0) = (2x, 2y, 2z)|_{(1,0,0)} = (2, 0, 0)$$

Entonces el plano tangente es

$$\nabla F(1, 0, 0) \cdot (x - 1, y - 0, z - 0) = 0$$

es decir,

$$x = 1$$

$\square$

Ejemplo 1.33. Hallar un vector normal a la superficie definida por:

$$2xy^3z + z \ln x + y \operatorname{sen} y = 0$$

en $(1, 2\pi, 0)$.

Solución. Ciertamente el punto $(1, 2\pi, 0)$ está en la superficie. Para encontrar un vector normal derivamos

$$\nabla F(x, y, z) = \left(2y^3z + \frac{z}{x}, 6xy^2z + y \cos y + \operatorname{sen} y, 2xy^3 + \ln x\right)$$

y evaluamos en $(1, 2\pi, 0)$ para obtener

$$\nabla F(1, 2\pi, 0) = (0, 2\pi, 2(2\pi)^3)$$

Como vimos, el gradiente $\nabla F(1, 2\pi, 0) = (0, 2\pi, 2(2\pi)^3)$ es normal a la superficie. $\square$

1.7.1. Caso del grafo de una función

En el caso que $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, de la definición 1.6 tenemos que el grafo de f corresponde a:

$$G(f) = \{(\boldsymbol{x}, f(\boldsymbol{x})) : \boldsymbol{x} \in A\} = \{(\boldsymbol{x}, z) : z = f(\boldsymbol{x})\}$$

Si definimos

$$F(\boldsymbol{x}, z) = z - f(\boldsymbol{x})$$

entonces las dos definiciones de plano tangente que hemos dado son coherentes. En efecto:

$$\nabla F(\boldsymbol{x}, z) = \left(-\frac{\partial f}{\partial x_1}, \ldots, -\frac{\partial f}{\partial x_n}, 1\right)$$

Entonces el plano tangente queda determinado por

$$\nabla F(\boldsymbol{x}, z_0)((\boldsymbol{x}, z) - (\boldsymbol{x}_0, z_0)) = 0$$

es decir,

$$-\nabla f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + (z - z_0) = 0$$

o sea, $z = z_0 + \nabla f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)$. Notar bien que el vector $(-\nabla f(\boldsymbol{x}_0), 1)$ es ortogonal al grafo de f en $(\boldsymbol{x}_0, f(\boldsymbol{x}_0))$

Ejemplo 1.34. Encontrar el plano tangente al grafo de la función:

$$f(x, y, z, t) = xe^{yt} + \cos zt$$

en $(3, 0, 0, 3)$.

Solución. Tenemos que $f(3, 0, 0, 3) = 3 + 1 = 4$. Además, derivando obtenemos

$$\nabla f(x, y, z, t) = (e^{yt}, xte^{yt}, -t \operatorname{sen}(zt), xye^{yt} - z \operatorname{sen}(zt))$$

y evaluando

$$\nabla f(3, 0, 0, 3) = (1, 9, 0, 0)$$

En consecuencia el plano tangente es

$$w - 4 = 1 \cdot (x - 3) + 9(y - 0) + 0(z - 0) + 0(t - 3)$$

y simplificando

$$w - 4 = x - 3 + 9y \qquad \text{ó} \qquad 9y + x - w = -1$$

□

Ejemplo 1.35. Hallar un vector unitario, normal a la superficie S dada por

$$z = x^2y^2 + y + 1$$

en $(0, 0, 1)$.

Solución. $F = z - x^2y^2 - y - 1 = 0$ y sus derivadas parciales son

$$\frac{\partial F}{\partial x} = -2xy^2 \ , \ \frac{\partial F}{\partial y} = -2x^2y - 1 \ , \ \frac{\partial F}{\partial z} = 1$$

Evaluando en $(0, 0, 1)$ obtenemos $\nabla F(0, 0, 1) = (0, -1, 1)$ y de aquí $\boldsymbol{v} = \frac{1}{\sqrt{2}}(0, -1, 1)$ es normal unitario. □

1.8. Ejercicios

Base algebraica y geométrica de $\mathbb{R}^m$

Ejercicio 1. Sean $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

1. Demuestre que $|\boldsymbol{x} \cdot \boldsymbol{y}| = \|\boldsymbol{x}\|\|\boldsymbol{y}\|$ si y slo si existe un escalar $\lambda \neq 0$ tal que $\boldsymbol{y} = \lambda \boldsymbol{x}$.
2. Deduzca que $\|\boldsymbol{x} + \boldsymbol{y}\| = \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ implica que $\boldsymbol{y} = \lambda \boldsymbol{x}$, con $\lambda > 0$.

Ejercicio 2. Sean $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. Demuestre que:

$$\frac{\|\boldsymbol{x} - \boldsymbol{z}\|}{1 + \|\boldsymbol{x} - \boldsymbol{z}\|} \leq \frac{\|\boldsymbol{x} - \boldsymbol{y}\|}{1 + \|\boldsymbol{x} - \boldsymbol{y}\|} + \frac{\|\boldsymbol{y} - \boldsymbol{z}\|}{1 + \|\boldsymbol{y} - \boldsymbol{z}\|}$$

Ejercicio 3. Demostrar las siguientes identidades:

1. **Identidad de Polarización:**
$$\boldsymbol{x} \cdot \boldsymbol{y} = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$
2. **Ley del paralelógramo:**
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Ejercicio 4. Se define la distancia de un punto $\boldsymbol{x} \in \mathbb{R}^n$ a un conjunto $A \subseteq \mathbb{R}^n$ como:

$$d_A(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in A} \|\boldsymbol{y} - \boldsymbol{x}\|$$

Demostrar que $\mathrm{fr}(A) = \{\boldsymbol{x} \in \mathbb{R}^n : d_A(\boldsymbol{x}) = 0 \ \wedge \ d_{A^c}(\boldsymbol{x}) = 0\}$.
Ind: Sea A un conjunto no vacío y acotado, entonces si $a = \inf(A)$ se tiene que $\forall \varepsilon > 0 \ \exists x \in A$ tal que $a + \varepsilon \geq x$.

Funciones con valores en $\mathbb{R}^m$

Ejercicio 5. Grafique:

1. Las superficies de nivel para la función de tres variables definida por $f(x, y, z) = x^2 + y^2 - z^2$.
2. Las curvas de nivel de la función $f(x, y) = x^2 - y^2$.
3. Las curvas de nivel de la función $f(x, y) = x^3 - 3xy^2$.
4. Las curvas de nivel y el grafo de la función $f(x, y) = x + y$.
5. Las curvas de nivel y el grafo de la función $f(x, y) = x^2 + 4y^2$.
6. El grafo de la función de dos variables definida por $f(x, y) = x^2 - y^2$.

Ejercicio 6. Encuentre los conjuntos de nivel para las siguientes funciones (para los niveles que se indican).

1. $f(x,y) = x + y$ para $f(x,y) = 1$.
2. $f(x,y) = (x^2 + y^2 + 1)^2 - 4x^2$ para $f(x,y) = 0$.
3. $f(x,y,z) = (xyz, x + y)$ para $f(x,y,z) = (0,1)$.

Ejercicio 7. Sea f definida por la siguiente fórmula:

$$f(x_1,x_2) = \begin{cases} \frac{x_1}{\sqrt{x_1^2 - x_2}} & \text{si } (x_1,x_2) \neq 0 \\ 0 & \text{si } (x_1,x_2) = 0 \end{cases}$$

1. Encuentre el conjunto donde se puede definir f, es decir $Dom(f)$. Grafique.
2. Determine las curvas de nivel de f.
3. Determine si f es continua en $(0,0)$.

Conceptos introductorios de topología

Ejercicio 8. Sea $A = \{(x,y) \in \mathbb{R}^2 : 0 < x < 2,\ 0 < y < 1\}$.

1. Calcule $\text{int}(A), \text{adh}(A), \text{der}(A)$ y $\text{fr}(A)$.
2. Deduzca que A es abierto.

Ejercicio 9. Sea $B = \{(x,y) \in \mathbb{R}^2 : x^2 < y < x + 1\}$.

1. Calcule $\text{int}(B), \text{adh}(B), \text{der}(B)$ y $\text{fr}(B)$.
2. Deduzca que B es abierto.

Ejercicio 10. Sea $C = \{(x, \text{sen}(x) + \cos(x)) \in \mathbb{R}^2 : 0 \leq x \leq 2\pi\}$.

1. Calcule $\text{int}(C), \text{adh}(C), \text{der}(C)$ y $\text{fr}(C)$.
2. Deduzca que C es cerrado.

Ejercicio 11. Sea $D = \{\left(1 + \frac{1}{n}, 1 - \frac{1}{n^2}\right) : n \in \mathbb{N} \setminus \{0\}\}$.

1. Calcule $\text{int}(D), \text{adh}(D), \text{der}(D)$ y $\text{fr}(D)$.
2. Deduzca que D no es abierto ni cerrado.

Ejercicio 12. Sea $E = \{\left(a + \frac{1}{n}, b + \frac{1}{m}\right) : m, n \in \mathbb{Z},\ p, q \in \mathbb{N} \setminus \{0\}\}$.

1. Calcule $\text{int}(E), \text{adh}(E), \text{der}(E)$ y $\text{fr}(E)$.
2. Es E abierto?
3. Es E cerrado?

Ejercicio 13. Demuestre las siguientes proposiciones:

1. A es abierto sí y sólo sí $\text{int}(A) = A$.
2. A es cerrado sí y sólo sí $\text{adh}(A) = A$.

Ejercicio 14.

1. Demuestre que la unión cualquiera de conjuntos abiertos es un conjunto abierto.
2. Demuestre que la intersección finita de conjuntos abiertos es un conjunto abierto.
3. Dé un ejemplo en que la intersección infinita de conjuntos abiertos no sea un conjunto abierto.
4. Demuestre que la intersección cualquiera de conjuntos cerrados es un conjunto cerrado.
5. Demuestre que la unión finita de conjuntos cerrados es un conjunto cerrado.
6. Dé un ejemplo en que la unión infinita de conjuntos cerrados no sea un conjunto cerrado.

Ejercicio 15. Sean A, B subconjuntos de $\mathbb{R}^n$. Demostrar que:

1. $A \subseteq B \ \Rightarrow \text{int}(A) \subseteq \text{int}(B) \ \wedge \ \text{adh}(A) \subseteq \text{adh}(B)$.
2. $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$. Dé un ejemplo en que no se cumpla la inclusión inversa.
3. $\text{int}(A) \cap \text{int}(B) = \text{int}(A \cap B)$.
4. $\text{adh}(A) \cup \text{adh}(B) = \text{adh}(A \cup B)$.
5. $\text{adh}(A \cap B) \subseteq \text{adh}(A) \cap \text{adh}(B)$. Dé un ejemplo en que no se cumpla la inclusión inversa.
6. $\text{int}(A^c) = (\text{adh}(A))^c$.
7. $\text{adh}(A^c) = (\text{int}(A))^c$.

Ejercicio 16. Sea $E = \{(x, \text{sen}(\frac{1}{x})) : x > 0\}$. Demuestre que:

$$\text{adh}(E) = E \cup \{(0, y) : -1 \leq y \leq 1\}.$$

Límites y continuidad

Ejercicio 17. Determine si las siguientes funciones admiten límite en los puntos que se indican

1. $f(x, y) = 2xy\frac{x^2-y^2}{x^2+y^2}$ en $(x_0, y_0) = (0, 0)$
2. $f(x, y) = \frac{\text{sen}(x)-\text{sen}(y)}{x-y}$ en $(x_0, y_0) = (0, 0)$
3. $f(x, y) = \frac{xy^2}{x^2+y^4}$ en $(x_0, y_0) = (0, 0)$

Ejercicio 18. Determine la existencia del límite de las siguientes funciones $f : \mathbb{R}^2 \to \mathbb{R}$ en $(0,0)$:

1. $\frac{xy^2}{x^2+y^2}$
2. $2xy\frac{x^2-y^2}{x^2+y^2}$
3. $\frac{xy^2}{x^2+y^4}$
4. $\frac{\tan(x)-\tan(y)}{\cot(x)-\cot(y)}$
5. $|x|\sqrt{\frac{x^2+y^2}{x^2-y^2}}$
6. $|x|\sqrt{\frac{x^2-y^2}{x^2+y^2}}$
7. $\frac{xy^2}{(x^2+y^2)\sqrt{x^2+y^2}}$
8. $\frac{\operatorname{sen}(xy)}{xy^3}$

Ejercicio 19. Sea $\lim_{x\to x_0} f(x) = 3$. Demuestre usando $\varepsilon - \delta$ que

$$\lim_{x\to x_0} f^3(x) = 27$$

Ejercicio 20. Demuestre la siguiente caracterización de funciones continuas en A:
$f : A \longrightarrow \mathbb{R}^m$ es continua en A sí y sólo sí la preimagen de todo abierto de $\mathbb{R}^m$ es la intersección de A con un abierto de $\mathbb{R}^n$.

Ejercicio 21. Determine los puntos crt́icos de la función

$$f(x,y,z) = x^3 - xz + yz - y^3 + 2z^3$$

En base a su resultado determine la continuidad de f.

Ejercicio 22. Sean $f, g, h : \mathbb{R}^2 \to \mathbb{R}$ tres funciones continuas. Se define la función $F : \mathbb{R}^2 \to \mathbb{R}$ por

$$F(x,y) = h(f(x,y), g(x,y))$$

Demuestre que F es continua.

Ejercicio 23. Determine la continuidad de las siguientes funciones:

1. $f(x,y) = \begin{cases} \frac{x}{x+y} & \text{si } x+y \neq 0 \\ 1 & \text{si } x+y = 0 \end{cases}$
2. $f(x,y,z) = x^3 \ln(x^3 y + z) + \operatorname{sen}(z^2 + x)$

Ejercicio 24. Para este ejercicio y el siguiente recordemos la siguiente propiedad: Sobre un e.v.n. E una función $f : E \to \mathbb{R}$ es continua en $\boldsymbol{x}_0$ si

$$\forall\varepsilon > 0\, \exists\delta > 0\, \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \Rightarrow |f(\boldsymbol{x}) - f(\boldsymbol{x}_0)| < \varepsilon \Leftrightarrow$$

$$\forall\varepsilon > 0\, \exists\delta > 0\, B(\boldsymbol{x}_0, \delta) \subseteq f^{-1}(B(f(\boldsymbol{x}_0, \varepsilon)))$$

Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una función lineal. Demuestre que las siguientes proposiciones son equivalentes:

1. L es continua en todo punto de $\mathbb{R}^n$.
2. L es continua en $\mathbf{0} \in \mathbb{R}^n$.
3. $\|L(x)\|$ es acotada si $x \in \overline{B}(0,1)$.

Ejercicio 25. Sea $L : E \to \mathbb{R}$ una función lineal. Demuestre que las siguientes proposiciones son equivalentes.

1. L es continua en E.
2. L es continua en $0 \in E$.
3. $\sup\limits_{\boldsymbol{x} \in E, \boldsymbol{x} \neq 0} \frac{|L\boldsymbol{x}|}{\|\boldsymbol{x}\|} = \sup\limits_{\|\boldsymbol{x}\| \leq 1} |L\boldsymbol{x}| < +\infty$

Ejercicio 26. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$. Pruebe que las siguientes proposiciones son equivalentes:

1. f es continua en todo punto de $\mathbb{R}^n$.
2. $(\forall A \subseteq \mathbb{R}^m$ abierto$)$ $f^{-1}(A)$ es abierto en $\mathbb{R}^n$.
3. $(\forall B \subseteq \mathbb{R}^m$ cerrado$)$ $f^{-1}(A)$ es cerrado en $\mathbb{R}^n$.

Indicación. Para la segunda pruebe antes que $f^{-1}(A^C) = f^{-1}(A)^C$.

Diferenciabilidad de funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$

Ejercicio 27. Según los teoremas vistos en la sección 1.6 tenemos las siguientes implicaciones:

1. Derivadas parciales continuas en $\boldsymbol{x}_0 \Rightarrow f$ diferenciable en $\boldsymbol{x}_0 \Rightarrow f$ es continua en $\boldsymbol{x}_0$.
2. f diferenciable en $\boldsymbol{x}_0 \Rightarrow$ existen todas las derivadas parciales en $\boldsymbol{x}_0$.

Encontrar ejemplos donde se muestra que las recíprocas de a) y b) son falsas.

Ejercicio 28. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ definida por $f(\boldsymbol{x}) = \boldsymbol{a} + A\boldsymbol{x}$, donde $\boldsymbol{a} \in \mathbb{R}^n$ y $A \in M_{m\times n}(\mathbb{R})$. Demuestre que:

1. f es continua en $\mathbb{R}^n$.
2. f es diferenciable en todo $\boldsymbol{x}$ y $Df(\boldsymbol{x}) = A$.

Ejercicio 29. Demuestre por definición que $f(x, y) = \operatorname{sen}(x + y)$ es diferenciable en $(0, 0)$.

Ejercicio 30. Considere las funciones $f : \mathbb{R} \times (-\pi/2, \pi/2) \to \mathbb{R}^2$ y $g : \mathbb{R}_{++} \times \mathbb{R} \to \mathbb{R}^2$ definidas por:

$$f(u, v) = (u\cos(v), u\ \operatorname{sen}(v))$$
$$g(x, y) = (\sqrt{x^2 + y^2}, artan(y/x))$$

1. Encontrar $D(g \circ f)(u, v)$ y $D(f \circ g)(x, y)$.
2. Determinar si $Dom(f) = Dom(g \circ f)$, y si $Dom(g) = Dom(g \circ f)$.

Ejercicio 31. Encuentre la matriz Jacobiana de la función definida por

$$f(x, y) = (xe^y, x^2 + y\cos(x + y), \tanh(xy))$$

Ejercicio 32. Sea $g : \mathbb{R}^3 \to \mathbb{R}^2$ y $f : \mathbb{R}^2 \to \mathbb{R}^2$ donde $g(1,1,1) = (2,3)$. Se tiene que

$$Df(2,3) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{y } f \circ g(x,y,z) = \begin{bmatrix} xy^2z^2 \\ x^2y^2z \end{bmatrix}$$

En base a esto calcule $\frac{\partial g_2}{\partial y}(1,1,1)$.

Ejercicio 33. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

1. $f(x,y) = \begin{cases} \frac{x}{x+y} & \text{si } x+y \neq 0 \\ 1 & \text{si } x+y = 0 \end{cases}$

2. $f(x,y) = \begin{cases} \frac{|x|}{y^2} \exp\left(-\frac{|x|}{y^2}\right) & \text{si } y \neq 0 \\ 0 & \text{si } y = 0 \end{cases}$

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

4. $f(x,y) = \begin{cases} \frac{xy^3}{x^3+y^3} & \text{si } x \neq -y \\ 0 & \text{si } x = -y \end{cases}$

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

6. $f(x,y) = \begin{cases} xy^2 \operatorname{sen}\left(\frac{1}{xy}\right) & \text{si } xy \neq 0 \\ k & \text{si } xy = 0 \end{cases}$

7. $f(x,y) = \begin{cases} y\frac{x^2-y^2}{x^2+y^2} & \text{si } (x,y) \neq 0 \\ k & \text{si } (x,y) = 0 \end{cases}$

8. $f(x,y) = \begin{cases} \frac{xy^k}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$

En cada uno de los casos:

1. Determine la continuidad de f y determine, en los casos que corresponda, las condiciones sobre los parametros para que se cumpla la continuidad.
2. Determine la diferenciabilidad de f y determine, en los casos que corresponda, las condiciones sobre los parametros para que se cumpla la diferenciabilidad. Calcule todas sus derivadas parciales (si existen).
3. Determine la continuidad de las derivadas parciales.
4. Calcule las segundas derivadas de f en los casos que sea posible y en caso de que las derivadas cruzadas no sean simétricas explique por qué.

Ejercicio 34. Sea $S = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| = 1\}$. Sea $g : S \to \mathbb{R}$ una función continua tal que

$$g(1,0) = g(0,1) = 0 \text{ y } g(-\boldsymbol{x}) = -g(\boldsymbol{x}), \forall \boldsymbol{x} \in \mathbb{R}^2$$

Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función definida por

$$f(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\| g\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right) & \text{si } \boldsymbol{x} \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

1. Dado $\boldsymbol{a} \in \mathbb{R}^2$ demuestre que la función $h(t) = f(t\boldsymbol{a}), t \in \mathbb{R}$ es diferenciable.
2. ¿Es f diferenciable en $(0,0)$?

Ejercicio 35. Sea $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ una función de clase $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R}^2)$, tal que sus componentes ϕ_1, y ϕ_2 verifican

$$\frac{\partial \phi_2}{\partial y} = \frac{\partial \phi_1}{\partial x} \text{ y } \frac{\partial \phi_2}{\partial x} = -\frac{\partial \phi_1}{\partial y}$$

Sea $h : \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathcal{C}^1(\mathbb{R}^2, \mathbb{R})$. Se define la función $f : \mathbb{R}^2 \to \mathbb{R}$ por $f = h \circ \phi$. Demuestre que

$$\langle \nabla f(x, y), \nabla \phi_1(x, y) \rangle = \frac{\partial h}{\partial u}(\phi(x, y)) \cdot \|\nabla \phi_1(x, y)\|^2.$$

Ejercicio 36. Una función $u = f(x, y)$ con segundas derivadas parciales continuas que satisfaga la ecuación de Laplace

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

se llama función armónica. Determinar cuales de las siguientes funciones son armónicas.

1. $u(x, y) = x^3 - 3xy^2$
2. $u(x, y) = \operatorname{sen}(x) \cosh(y)$
3. $u(x, y) = e^x \operatorname{sen}(y)$.

Ejercicio 37. Si $g(x, y) = e^{x+y}$, $f'(0) = (1, 2)$, encontrar $F'(0)$ donde

$$F(t) = g(f(t)) \ (f : \mathbb{R} \to \mathbb{R}^2) \qquad \text{y} \qquad f(0) = (1, -1)$$

Ejercicio 38. Si $f(x, y, z) = sen(x), F(t) = (cos(t), sen(t), t)$, encontrar $g'(\pi)$ donde $g(t) = f(F(t))$.

Gradiente y geometría

Ejercicio 39. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ continua y diferenciable. Sea $G_f = \{[(x, y), f(x, y)] : (x, y) \in Dom(f)\}$ el grafo de f. Sea $F : \mathbb{R}^3 \to \mathbb{R}$ tal que $F(x, y, z) = z - f(x, y)$.

1. Muestre que G_f corresponde a un conjunto de nivel de F.
2. Demuestre que $\nabla F = (-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1)$.
3. Encuentre el vector normal y el plano tangente a G_f cuando $f(x, y) = x + ye^x$ en el punto $(1, 1)$.

Ejercicio 40. Encontrar el gradiente ∇f en cada uno de los siguientes casos.

1. $f(x, y) = \log_x y$
2. $f(x, y) = x^2 - y^2 \operatorname{sen}(y)$ en (a, b)
3. $f(x, y, z) = x^2 + y^2 - z^2$
4. $f(x, y) = (e^x \cos(y), e^x \operatorname{sen}(y))$
5. $f(x, y, z) = (x - \ln(x), e^{\operatorname{sen}(y)}, |z^3|)$
6. $f(x, y, z) = (5x^4 - z^2, 2yz^3 + 3y^2, 3y^2z^2)$
7. $f(x, y, z) = (x^2 + y^2, xyz, z^{\ln(x)})$

Ejercicio 41. Encontrar la derivada direccional para las siguientes funciones en la dirección u que se indica y en el punto a dado:

1. $f(x,y,z) = xyz$, $v = (\cos(\alpha)\operatorname{sen}(\beta), \operatorname{sen}(\alpha)\operatorname{sen}(\beta), \cos(\beta))$ y $a = (1,0,0)$.
2. $f(x,y) = e^x \operatorname{sen}(y)$, $u = (\cos(\alpha), \operatorname{sen}(\alpha))$ y $a = (1,0)$.
3. $f(x,y,mz) = x^2 + ye^z$, u dirección unitaria determinada por la recta tangente a la curva $g(t) = (3t^2 + t + 1, 2^t, t^2)$ en $g(0)$ y $a = (1,1,0)$.

Ejercicio 42. Encuentre 3 casos en los que una función $f : \mathbb{R}^3 \to \mathbb{R}$ tenga derivadas direccionales en un punto $\boldsymbol{x}_0$, sin ser diferenciable en $\boldsymbol{x}_0$.

Ejercicio 43. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función definida por

$$f(x,y) = \begin{cases} (x^2+y^2)\operatorname{sen}\left(\frac{1}{(x^2+y^2)^{1/2}}\right) & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

1. Calcular $\nabla f(0,0)$.
2. Probar que f es diferenciable en $(0,0)$.
3. Probar que $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ no son continuas en $(0,0)$.

Ejercicio 44. Sean f y g funciones de $\mathbb{R}^3 \to \mathbb{R}$. Suponer que f es diferenciable y $\nabla f(x) = g(x) \cdot x$. Mostrar que f es constante para las esferas centradas en el origen.

Ejercicio 45. Hallar la ecuación para el plano tangente a cada superficie $z = f(x,y)$ en el punto indicado:

1. $z = x^3 + y^3 - 6xy$ en $(1,2,-3)$.
2. $z = \cos(x)\operatorname{sen}(y)$ en $(0, \pi/2, 1)$

Ejercicio 46. Calcular para los siguientes casos la dirección de mayor crecimiento en $(1,1,1)$.

1. $f(x,y,z) = xy + yz + xz$
2. $f(x,y,z) = \frac{1}{x^2+y^2+z^2}$

Ejercicio 47. Sean $f(x,y) = x^2+y^2$ y $g(x,y) = e^{xy}$. Determine los puntos (x,y) donde la derivada direccional de f en la dirección de máximo crecimiento de g es igual a la derivada direccional de g en la dirección de máximo crecimiento de f.

Ejercicio 48. Considere las funciones

$$f(x,y) = (\tan(x+y), 1+xy, e^{x^2+y}) \quad \text{y} \quad g(u,v,w) = \operatorname{sen}(uv + \pi w)$$

Sea $h = g \circ f$. Calcule la ecuación del plano tangente a h en $(0,0,0)$

CAPÍTULO 2

Derivadas de Orden Superior

En este capítulo estudiaremos las segundas derivadas y derivadas enésimas de funciones de varias variables para luego extendernos con una generalización de la aproximación de Taylor. Ambos conceptos nos permitirán ver de forma introductoria la convexidad de funciones y criterios de primer y segundo orden para la optimización con y sin restricciones.

2.1. Derivadas de orden superior y teorema de Taylor

Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en A, entonces

$$\frac{\partial f}{\partial x_i} : A \subset \mathbb{R}^n \to \mathbb{R}$$

define una función. Como tal esta función puede tener derivadas parciales y también ser diferenciable. Cuando $\frac{\partial f}{\partial x_i}$ tiene derivada parcial con respecto a x_j en $\boldsymbol{x}_0$, la anotamos como

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{x_0}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x_0})$$

Ejemplo 2.1. Calcular $\frac{\partial^2 f}{\partial x \partial y}$ y $\frac{\partial^2 f}{\partial x^2}$ si $f(x,y) = ln(x^2+y^2)$.

Solución.

$$\frac{\partial f}{\partial x} = \frac{2x}{x^2+y^2}$$

Entonces

$$\frac{\partial^2 f}{\partial x^2} = \frac{2(x^2+y^2) - 2x(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

y

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{-2x(2y)}{(x^2+y^2)^2}$$

También podemos calcular

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{-2y(2x)}{(x^2+y^2)^2}$$

Notemos que en este caso $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. Esto no es sólo coincidencia como veremos más adelante. Además, podemos observar que la función f satisface

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

Por esta razón f se dice armónica. La combinación

$$\triangle f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

se conoce como el Laplaciano de f. $\square$

Definición 2.1. Se define el laplaciano de $f : \mathbb{R}^n \to \mathbb{R}$ como:

$$\triangle f(\boldsymbol{x}) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(\boldsymbol{x})$$

Para $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Recordemos que f se dice de clase $\mathcal{C}^1$ en un dominio A si f es diferenciable en A y todas sus derivadas parciales son continuas en A.

Definición 2.2. f se dirá dos veces diferenciables en $\boldsymbol{x}_0$ si f es diferenciable en una vecindad de $\boldsymbol{x}_0$ y todas sus derivadas parciales son diferenciables en $\boldsymbol{x}_0$. f se dirá de clase $\mathcal{C}^2$ en A si todas sus segundas derivadas parciales son continuas en A.

Es fácil extender estas definiciones a órdenes superiores.

A continuación veremos uno de los resultados más importantes de esta sección que dice relación con la posibilidad de intercambiar el orden de las derivadas

Teorema 2.1. (Teorema de Schwarz)
Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función dos veces diferenciable en A. Si las segundas derivadas parciales son continuas en $\boldsymbol{x}_0 \in A$ entonces:

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0) \qquad \forall i, j = 1, \ldots, n$$

Demostración. Una pequeña reflexión lleva a concluir que basta estudiar el caso en que $n = 2$. Dado $\boldsymbol{x}_0 = (x_{01}, x_{02}) \in A$ y $\boldsymbol{h} = (h_1, h_2) \in \mathbb{R}^2$, $\|\boldsymbol{h}\| < r$, definimos $F : B(\boldsymbol{0}, r) \to \mathbb{R}$ de la siguiente forma

$$F(h_1, h_2) = [f(x_{01} + h_1, x_{02} + h_2) - f(x_{01} + h_1, x_{02})] - [f(x_{01}, x_{02} + h_2) - f(x_{01}, x_{02})]$$

Notemos que F queda bien definida para r pequeño. Sea $g(t) = f(t, x_{02} + h_2) - f(t, x_{02})$. Entonces g es diferenciable y por el teorema del valor medio (teorema 1.23) existe h_1' tal que $|h_1'| < |h_1| \leq \|\boldsymbol{h}\|$ y

$$F(h_1, h_2) = g(x_{01} + h_1) - g(x_{01}) = g'(x_{01} + h_1')h_1$$

calulando g' y evaluando, se obtiene que

$$F(h_1, h_2) = \left[\frac{\partial f}{\partial x}(x_{01} + h_1', x_{02} + h_2) - \frac{\partial f}{\partial x}(x_{01} + h_1', x_{02})\right] h_1$$

Sea ahora $\tilde{g}(t) = \frac{\partial f}{\partial x}(x_{01} + h_1', t)$. Entonces $\tilde{g}$ es diferenciable y nuevamente por el teorema del valor medio se tiene que existe h_2' tal que $|h_2'| < |h_2| \leq \|\boldsymbol{h}\|$ y además:

$$F(h_1, h_2) = (\tilde{g}(x_{02} + h_2) - \tilde{g}(x_{02}))h_1 = \tilde{g}'(x_{02} + h_2')h_2 h_1$$

Calculando $\tilde{g}'$ y evaluando se obtiene que

$$\frac{F(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial y \partial x}(x_{01} + h_1', x_{02} + h_2')$$

Pero uno también puede escribir F de la siguiente manera

$$F(h_1, h_2) =$$
$$[f(x_{01} + h_1, x_{02} + h_2) - f(x_{01}, x_{02} + h_2)] - [f(x_{01} + h_1, x_{02}) - f(x_{01}, x_{02})]$$

y podemos repetir las mismas aplicaciones del teorema del valor medio para probar que

$$\frac{F(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial x \partial y}(x_{01} + h_1'', x_{02} + h_2'')$$

con $|h_1''| < |h_1|$ y $|h_2''| < |h_2|$. Por lo tanto

$$\frac{\partial^2 f}{\partial y \partial x}(x_{01} + h_1', x_{02} + h_2') = \frac{\partial^2 f}{\partial x \partial y}(x_{01} + h_1'', x_{02} + h_2'')$$

y finalmente, haciendo $h_1 \to 0$ y $h_2 \to 0$, por la continuidad de las derivadas parciales se obtiene que

$$\frac{\partial^2 f}{\partial x \partial y}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{x}_0)$$

■

Ejemplo 2.2. El siguiente ejemplo nos muestra que las derivadas cruzadas pueden ser distintas cuando no se cumplen las hipótesis del teorema. Consideremos la función

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

En todo punto $(x, y) \neq (0, 0)$ derivamos directamente y se obtiene

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}$$

en $(x, y) = (0, 0)$ calculamos por definición la derivada parcial

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0$$

Por lo tanto, tenemos que

$$\frac{\partial f}{\partial x} = \begin{cases} \frac{y(x^4 + 4y^2x^2 - y^4)}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

Y calculando por definición nuevamente, se obtiene que

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y\to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y} = \lim_{y\to 0} \frac{-\frac{y^5}{y^4}}{y} = -1$$

Por otro lado, y de manera análoga, se tiene que

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

y de aquí se obtiene que

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1$$

Es decir,

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$$

¿Cómo se interpreta esto a la luz del teorema? Estudiemos la continuidad de las derivadas cruzadas, para $(x,y) \neq (0,0)$ derivando directamente se obtiene que

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2+y^3)^3}$$

Tomemos una sucesión de la forma $(x_n, 0) \to (0,0)$ y notemos que

$$\lim_{x_n\to 0} \frac{\partial^2 f}{\partial y \partial x}(x_n, 0) = \lim_{x_n \to 0} \frac{{x_n}^6}{{x_n}^6} = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0,0)$$

por lo tanto la función $\frac{\partial^2 f}{\partial y \partial x}$ es discontinua en $(0,0)$ $\square$

Corolario 2.1. *Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^k$ con $k \in \mathbb{N}$. Entonces:*

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} = \frac{\partial^k f}{\partial x_{\sigma(i_1)} \partial x_{\sigma(i_2)} \ldots \partial x_{\sigma(i_k)}}$$

Donde $i_1, i_2, \ldots i_k \in \{1, \ldots, n\}$ y $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ es una permutación cualquiera.

Demostración. Basta aplicar el teorema de intercambio de derivadas reiterativamente. ■

Ahora que ya conocemos las derivadas de mayor orden y sus principales propiedades estamos preparados para estudiar los desarrollos de Taylor en varias variables. Recordemos el caso de una variable, en el cual vamos a basar nuestro argumento para varias variables.

Teorema 2.2. (Teorema de Taylor de 1^{er} orden)
Sea $f : (a,b) \to \mathbb{R}$ derivable en $x_0 \in (a,b)$ entonces es posible obtener una expresión equivalente a $f(x)$ dada por

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R_1(x_0, x - x_0)$$

donde el término $R_1(x_0, x - x_0)$ se denomina resto y satisface

$$\lim_{x \to x_0} \frac{R_1(x_0, x - x_0)}{|x - x_0|} = 0$$

que además puede ser descrito en términos de una integral si f es una función $\mathcal{C}^2$.

Demostración. Notemos que definiendo

$$R_1(x_0, x - x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0)$$

Por definición de $f'(x_0)$ se obtiene que

$$\lim_{x \to x_0} \frac{R_1(x_0, x - x_0)}{|x - x_0|} = \lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{|x - x_0|} = 0$$

Ahora, del 2^{do} teorema fundamental del cálculo (teorema 1.9) tenemos

$$f(x) - f(x_0) = \int_{x_0}^{x} f'(t)dt$$

entonces

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt$$

Si suponemos que f es de clase $\mathcal{C}^2$, integrando por partes $(u = f'(t), v = t - x)$ obtenemos

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x} (x - t)f''(t)dt$$

Así obtenemos la expresión integral para R_1 dada por

$$R_1(x, x_0) = \int_{x_0}^{x} (x - t)f''(t)dt$$

■

En general, si f es k veces derivable en x_0, entonces

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \ldots + \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k + R_k(x, x_0)$$

y el resto satisface

$$\lim_{x \to x_0} \frac{R_k(x, x_0)}{|x - x_0|^k} = 0$$

Si se supone además que f es de clase $\mathcal{C}^{k+1}$ en una vecindad de x_0, integrando por partes reiterativamente se obtiene la expresión integral para R_k dada por

$$R_k(x, x_0) = \int_{x_0}^{x} \frac{(x - t)^k}{k!} f^{(k+1)}(t)dt$$

Luego, si consideramos

$$M = \max_{t \in [x_0 - \delta, x_0 + \delta]} |f^{(k+1)}(t)| \text{ para } 0 < \delta < |x|$$

se tiene que

$$|R_k(x, x_0)| \leq M \left| \int_{x_0}^{x} \frac{(x - t)^k}{k!} dt \right| = \frac{M}{(k + 1)!} |x - x_0|^{k+1}$$

Nota 2.1. De lo anterior se deduce que todo polinomio es de clase $\mathcal{C}^\infty$, es decir, es infinitamente diferenciable y sus enésimas derivadas parciales son continuas.

Cuando $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ también podemos establecer un teorema de Taylor. Notamos que en caso que f tome valores en $\mathbb{R}^m$ lo que diremos se aplica a cada una de las funciones coordenada. Comenzamos con una definición.

Definición 2.3. Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en $\boldsymbol{x}_0 \in A$. Se define entonces la matriz Hessiana de f en $\boldsymbol{x}_0$ como

$$Hf(\boldsymbol{x_0}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x_0}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x_0}) \end{bmatrix}$$

Nota 2.2. Notamos que, por el teorema 2.1, si f es de clase $\mathcal{C}^2$ en una vecindad de $\boldsymbol{x}_0$ entonces la matriz Hessiana de f en $\boldsymbol{x}_0$ es simétrica.

Teorema 2.3. (Teorema de Taylor de 2^{do} orden)
Sea $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^3$. Entonces

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\boldsymbol{h} + \frac{1}{2}\boldsymbol{h}^T Hf(\boldsymbol{x}_0)\boldsymbol{h} + R_2(\boldsymbol{x}_0, \boldsymbol{h})$$

donde el resto $R_2(\boldsymbol{x}_0, \boldsymbol{h})$ tiene una expresión integral y $|R_2(\boldsymbol{x}_0, \boldsymbol{h})| \leq M\|\boldsymbol{h}\|^3$. Así que en particular

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{R_2(\boldsymbol{x}_0, \boldsymbol{h})}{\|\boldsymbol{h}\|^2} = 0$$

Nota 2.3. Se puede demostrar que si f es de clase $\mathcal{C}^2$ entonces también se tiene que

$$\lim_{\boldsymbol{x} \to \boldsymbol{0}} \frac{R_2(\boldsymbol{x}_0, \boldsymbol{h})}{\|\boldsymbol{h}\|^2} = 0$$

aunque la expresión integral del resto no se tiene necesariamente.

Demostración. Utilizando la regla de la cadena tenemos

$$\frac{d}{dt} f(\boldsymbol{x}_0 + t\boldsymbol{h}) = Df(\boldsymbol{x}_0 + t\boldsymbol{h})\boldsymbol{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i$$

e integrando entre 0 y 1

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) - f(\boldsymbol{x}_0) = \sum_{i=1}^{n} \int_0^1 \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i dt$$

Utilizando nuevamente la regla de la cadena obtenemos para $u = \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i$

$$\frac{du}{dt} = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_j h_i$$

Integrando por partes, considerando u como arriba y $v = t - 1$

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)h_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i h_j dt$$

que es un resultado correspondiente a la fórmula de Taylor de primer orden (teorema 2.2)

$$f(\boldsymbol{x}_0+\boldsymbol{h}) = f(\boldsymbol{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)h_i + R_1(\boldsymbol{x}_0,\boldsymbol{h})$$

Integraremos nuevamente por partes, con $v = -(t-1)^2/2$ y $u = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i h_j$. Notemos que

$$\frac{du}{dt} = \sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i h_j h_k$$

Por lo tanto obtenemos

$$f(\boldsymbol{x}_0+\boldsymbol{h}) = f(\boldsymbol{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)h_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{2}\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0)h_i h_j + R_2(\boldsymbol{x}_0,\boldsymbol{h})$$

donde el resto tiene la forma integral

$$R_2(\boldsymbol{x}_0,\boldsymbol{h}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \int_0^1 \frac{(1-t)^2}{2} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i h_j h_k dt$$

Como $\frac{\partial f^3}{\partial x_k \partial x_j \partial x_i}$ son continuas en $\boldsymbol{x}_0$, existe $\delta > 0$ y $M > 0$ tal que

$$\|\boldsymbol{h}\| < \delta \;\Rightarrow\; \left| \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h}) \right| \leq M$$

La constante M puede tomarse como

$$M = \max_{\boldsymbol{x} \in B(\boldsymbol{x}_0,\delta)} \left| \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{x}) \right|$$

Por lo que, para todo $t \in [0,1]$ tenemos

$$\left| \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{x}_0 + t\boldsymbol{h})h_i h_j h_k \right| \leq M$$

Y finalmente

$$|R_2(\boldsymbol{x}_0,\boldsymbol{h})| \leq \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{1}{3!} M \|\boldsymbol{h}\|^3 = \left(\frac{n^3}{3!} M \right) \|\boldsymbol{h}\|^3$$

■

Ejemplo 2.3. Encontrar la fórmula de Taylor de orden 2 en torno a $x_0 = (0,0)$ para

$$f(x,y) = \operatorname{sen}(x+2y) + x^2$$

Solución. Evaluamos $f(0,0) = 0$. Después calculamos las derivadas parciales

$$\frac{\partial f}{\partial x}(x,y) = \cos(x+2y) + 2x \text{ y } \frac{\partial f}{\partial y}(x,y) = 2\cos(x+2y)$$

y evaluamos $\nabla f(0,0) = (1,2)$. Ahora calculamos las derivadas de segundo orden

$$\frac{\partial^2 f}{\partial x^2} = -\operatorname{sen}(x+2y) + 2,\; \frac{\partial^2 f}{\partial y \partial x} = -2\operatorname{sen}(x+2y) \text{ y } \frac{\partial^2 f}{\partial y^2} = -4\operatorname{sen}(x+2y)$$

y evaluamos

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2 \ , \ \frac{\partial^2 f}{\partial y \partial x}(0,0) = 0 \text{ y } \frac{\partial^2 f}{\partial y^2}(0,0) = 0$$

Considerando $\boldsymbol{h} = (h_1, h_2)$ tenemos finalmente

$$\begin{aligned} f(\boldsymbol{x}_0 + \boldsymbol{h}) &= f(h_1, h_2) = f(0,0) + (1,2)(h_1, h_2) + \frac{1}{2}(h_1, h_2) \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + R_2 \\ &= h_1 + 2h_2 + h_1^2 + R_2((0,0), \boldsymbol{h}) \end{aligned}$$

□

Continuando con este ejemplo, sabemos que la función $P_2(h_1, h_2) = h_1 + 2h_2 + h_1^2$ aproxima a $f(h_1, h_2)$ al segundo orden, en el sentido que $\frac{R_2((0,0),h)}{\|h\|^2} \Rightarrow 0$. Pero uno podría hacer una pregunta más específica: Si $\|h\| \leq 1/4$ ¿Qué error se comete por cambiar f y P_2? Tenemos la fórmula integral del resto o error

$$R_2(h_1, h_2) = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} \int_0^1 \frac{(1-t)^2}{2} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(x_0 + th) h_i h_j h_k dt.$$

Usando el teorema del valor medio integral (teorema 1.10) vemos que existen $t_{ijk} \in [0,1]$ tales que

$$R_2(h_1, h_2) = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} \frac{(1-t_{ijk})^2}{2} \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i}(x_0 + t_{ijk} h) h_i h_j h_k,$$

de donde podemos hacer nuestra estimación. En nuestro caso calculemos las terceras derivadas

$$\begin{array}{rcl} \dfrac{\partial^3 f}{\partial x^3} = -\cos(x+2y) & , & \dfrac{\partial^3 f}{\partial y^3} = -8\cos(x+2y) \\ \dfrac{\partial^3 f}{\partial x \partial y^2} = -4\cos(x+2y) & , & \dfrac{\partial^3 f}{\partial y \partial x^2} = -2\cos(x+2y) \end{array}$$

Luego, notando que $|\cos(x+2y)| \leq 1$ para cualquiera que sea el punto $(x,y) = x_0 + t_{ijk} h$, que $(1-t_{ijk})^2 \leq 1$ para cualquier $t_{ijk} \in [0,1]$, y que $|h_i| \leq \|\boldsymbol{h}\|$ para $i = 1, 2$. Se obtiene que

$$\begin{aligned} |R_2(h_1, h_2)| &\leq \frac{1}{2}(h_1^3 + 3 \cdot 2h_1^2 h_2 + 3 \cdot 4 h_1 h_2^2 + 8h_2^3) \\ &\leq \frac{1}{2}(1 + 6 + 12 + 8)\|h\|^3 = \frac{27}{2}\|\boldsymbol{h}\|^3 \end{aligned}$$

Así, si $\|\boldsymbol{h}\| \leq 1/4$ entonces $|R_2(h_1, h_2)| \leq 27/(2 \cdot 4^3)$

2.2. Extremos de funciones con valores reales

2.2.1. Condiciones de primer orden

Dentro de los puntos que pertenecen al dominio de la función, aquellos en donde esta alcanza un mínimo o un máximo revisten un especial interés por su importancia en muchos problemas prácticos.

Definición 2.4. Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A un subconjunto abierto de $\mathbb{R}^n$. Un punto $\boldsymbol{x}_0 \in A$ se dirá mínimo (máximo) local de f si existe una vecindad V de $\boldsymbol{x}_0$ tal que

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)\ \forall \boldsymbol{x} \in V \qquad (\text{resp. } f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0))$$

Un punto se dirá extremo local si es un mínimo o un máximo local. Un punto $\boldsymbol{x}_0$ se dirá crítico de f si $Df(\boldsymbol{x}_0) = 0$. Un punto crítico que no es extremo se dirá punto silla.

Teorema 2.4. *Sea $f : A \Rightarrow \mathbb{R}$ diferenciable, con A un subconjunto abierto de $\mathbb{R}^n$ y $\boldsymbol{x}_0 \in A$ un extremo local, entonces $\nabla f(\boldsymbol{x}_0) = 0$.*

Demostración. Si f tiene un mínimo local $\boldsymbol{x}_0$ definamos la función de una variable

$$g(t) = f(\boldsymbol{x}_0 + t\boldsymbol{h})$$

donde $\boldsymbol{h} \in \mathbb{R}^n$ es un punto cualquiera. Se tiene que g tiene un mínimo local en $t = 0$ pues

$$g(t) = f(\boldsymbol{x}_0 + t\boldsymbol{h}) \geq f\left(\boldsymbol{x}_0\right) = g\left(0\right)$$

y por tanto $g'(0) = 0$ lo que es equivalente a

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{h} = 0$$

y como esto es para cualquier $\boldsymbol{h} \in \mathbb{R}^n$, necesariamente $\nabla f(\boldsymbol{x}_0) = 0$. ■

Observemos que $\nabla f(\boldsymbol{x}_0) = 0$ es equivalente a:

$$\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0) = 0, \frac{\partial f}{\partial x_2}(\boldsymbol{x}_0) = 0, ..., \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0) = 0$$

en otras palabras, los puntos críticos de f satisfacen el sistema de ecuaciones

$$Df(\boldsymbol{x}) = 0$$

Este es un sistema de n ecuaciones con n incógnitas, pero en general no es un sistema lineal.

Ejemplo 2.4. Busque los puntos críticos de la siguiente función y clasifíquelos:

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

Solución. El sistema de ecuaciones que se obtiene es:

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xy + y^2 = 0 \\ \frac{\partial f}{\partial y} &= 2xy + x^2 = 0 \end{aligned}$$

de donde se tiene

$$\begin{aligned} y(2x + y) = 0 &\quad \Rightarrow \quad y = 0 \lor y = -2x \\ x(2y + x) = 0 &\quad \Rightarrow \quad x = 0 \lor x = -2y \end{aligned}$$

de la primera relación vemos que los puntos críticos deben ser de la forma $(a, 0),\ a \in \mathbb{R}$ o $(b, -2b),\ b \in \mathbb{R}$ mientras que de la segunda relación se observa que los puntos críticos deben ser de la forma $(0, c),\ c \in \mathbb{R}$ o $(-2d, d),\ d \in \mathbb{R}$. Como se tienen que cumplir ambas relaciones a la vez concluímos que el único punto crítico es $(0, 0)$. Por otro lado, notemos que en la recta $y = x$ la función vale $f(x, x) = 2x^3$ que toma valores positivos y negativos, como $f(0, 0) = 0$, concluímos que $(0, 0)$ no es ni máximo ni mínimo local, por lo tanto es un punto silla. □

2.2.2. Condiciones de segundo orden

Veamos ahora condiciones necesarias de 2^{do} orden equivalentes a las vistas en el caso de funciones de una variable, es decir, $f''(x) \geq 0$ para mínimo y $f''(x) \leq 0$ para máximo, y también condiciones suficientes equivalentes a las del caso de una variable, osea $f''(x) > 0$ para mínimo y $f''(x) < 0$ para máximo.

Definición 2.5. Sea A un subconjunto abierto de $\mathbb{R}^n$ y $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una función con derivadas de 2^{do} orden

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0)$$

en $\boldsymbol{x}_0$. El Hessiano de f en $\boldsymbol{x}_0$ es la función cuadrática definida por

$$Hf(\boldsymbol{x}_0)(\boldsymbol{d}) = \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0) d_i d_j$$

Otra forma de escribir la expresión anterior es

$$Hf(\boldsymbol{x}_0)(\boldsymbol{d}) = \frac{1}{2}\boldsymbol{d}^T H \boldsymbol{d}$$

donde

$$H = \begin{pmatrix} \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_1^2} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_1 \partial x_3} & \dots & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_2 \partial x_n} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_2^2} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_2 \partial x_3} & \dots & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & & \vdots \\ \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_n \partial x_2} & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_n \partial x_3} & \dots & \frac{\partial^2 f(\boldsymbol{x_0})}{\partial x_n^2} \end{pmatrix}$$

Ya habíamos introducido la matriz H, que se denomina matriz Hessiana de la función f. Recordemos que cuando f es de clase $\mathcal{C}^1$ esta matriz es simétrica ya que

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

Antes de dar el criterio de 2^{do} orden, recordemos algunos elementos de álgebra lineal.

Si A es una matriz simétrica, entonces A es diagonalizable, es decir, existe una matriz P tal que

$$A = PDP^T \tag{2.1}$$

donde P es una matriz invertible (ortogonal) y D es una matriz diagonal, es decir:

$$P^T P = I$$

donde I es la matriz identidad y

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

De (2.1) tenemos que

$$AP = PD$$

y esta ecuación por columnas es

$$Ap_i = \lambda_i p_i$$

donde p_i representa la i-ésima columna de P. Los λ_i son los valores propios de la matriz A mientras que los p_i son los vectores propios correspondientes a dichos valores propios.

Una matriz A se dice definida positiva si

$$\boldsymbol{x}^T A\boldsymbol{x} > 0 \ \forall \boldsymbol{x} \neq \boldsymbol{0}$$

Si además es simétrica, usando la diagonalización vista anteriormente, tenemos que

$$\boldsymbol{x}^T PDP^T \boldsymbol{x} > 0 \ \forall \boldsymbol{x} \neq \boldsymbol{0}$$

o sea (haciendo $P^T\boldsymbol{x} = \boldsymbol{y}$)

$$\boldsymbol{y}^T D\boldsymbol{y} > 0 \ \forall \boldsymbol{y} \neq \boldsymbol{0}$$

$$\Rightarrow \sum \lambda_i y_i^2 > 0 \ \forall \boldsymbol{y} \neq \boldsymbol{0}$$

y de aqui se obtiene el siguiente teorema.

Teorema 2.5. *Una matriz A es definida positiva si y sólo si los valores propios de A son positivos.*

Demostración. Haremos la demostración en dos partes. Sean

1. A es definida positiva ($\boldsymbol{x}^T A\boldsymbol{x} > 0, \ \forall \boldsymbol{x} \neq \boldsymbol{0}$).
2. Los valores propios de A son positivos.

(1) $\Rightarrow$ (2):
Sea λ un valor propio de A e $\boldsymbol{y} \neq \boldsymbol{0}$ un vector propio correspondiente al valor propio λ. Entonces

$$0 < \boldsymbol{y}^T A\boldsymbol{y} = \boldsymbol{y}^T(\lambda\boldsymbol{y}) = \lambda\boldsymbol{y}^T\boldsymbol{y} = \lambda\|\boldsymbol{y}\|^2 \ \Rightarrow \ \lambda > 0$$

(2) $\Rightarrow$ (1):
Suponiendo que A es simétrica entonces $A = PDP^T$, donde las columnas de P son una base ortonormal del vectores propios y D es la diagonal de los valores propios respectivos. De esta forma

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{x}^T PDP^T\boldsymbol{x} = (P^T\boldsymbol{x})^T DP^T\boldsymbol{x}$$

entonces definiendo $\boldsymbol{z} = P^T\boldsymbol{x}$ se tendrá que en términos de estas nuevas variables la forma cuadrática queda

$$\boldsymbol{x}^T A\boldsymbol{x} = \boldsymbol{z}^T D\boldsymbol{z} = \lambda_1 z_1^2 + \ldots + \lambda_n z_n^2 \geq 0$$

dado que los valores propios son positivos. La única forma de que la forma cuadrática sea nula es con $z_1 = \ldots = z_n = 0$, es decir $\boldsymbol{z} = \boldsymbol{0}$, pero entonces $P^T\boldsymbol{x} = 0 \ \Rightarrow \ \boldsymbol{x} = \boldsymbol{0}$ (se debe tener presente que $P^{-1} = P^T$). ■

Corolario 2.2. *Si A es definida positiva entonces existe $c > 0$ tal que*

$$\boldsymbol{x}^T A\boldsymbol{x} \geq c\,\|\boldsymbol{x}\|^2 \ \ \forall \boldsymbol{x} \in \mathbb{R}^n$$

donde, $c = \min\{\lambda_i : i = 1, ..., n\}$.

Demostración.

$$\begin{aligned}
\boldsymbol{x}^T A\boldsymbol{x} &= \sum \lambda_i y_i^2 \\
&\geq \text{mín}\{\lambda_i : i = 1, ..., n\} \|y\|^2 \\
&= c \left\|P^T \boldsymbol{x}\right\|^2 \\
&= c\boldsymbol{x}^T P P^T \boldsymbol{x} \text{ , recordemos que } PP^T = P^T P = I \\
&= c \|\boldsymbol{x}\|^2
\end{aligned}$$

■

También se habla de matriz semi-definida positiva cuando se satisface

$$\boldsymbol{x}^T A\boldsymbol{x} \geq 0 \ \forall \boldsymbol{x} \in \mathbb{R}^n$$

Teorema 2.6. *Una matriz A es semi-definida positiva si y solo si los valores propios de A son positivos o nulos.*

Nota 2.4. De forma análoga a las definiciones de matrices definidas y semi-definidas positivas, haciendo los cambios de desigualdad correspondientes, se obtienen las definiciones de matrices definidas y semi-definidas negativas.

Existen muchos criterios para determinar si una matriz es definida positiva o no, uno de los usados por su fácil comprobación es el siguiente: Una matriz B cuadrada $(n \times n)$ es definida positiva si y solo si todas las submatrices cuadradas a lo largo de la diagonal tienen determinantes positivos. Para el caso de las matrices definidas negativas los signos de los determinantes deben alternarse, comenzando con negativo.

De esta forma, cuando la matriz es semi-definida positiva se tiene que

$$|H_i| \geq 0 \ \forall i \in \{1, \ldots, j\}$$

mientras que cuando la matriz es semi-definida negativa se tiene que $|H_1| \leq 0, |H_2| \geq 0, \ldots$ tal que

$$(-1)^i |H_i| \geq 0 \ \forall i \in \{1, \ldots, j\}$$

Ejemplo 2.5. Para el caso de una función de tres variables los determinantes de las tres submatrices del Hessiano corresponden a

$$|H_1| = \frac{\partial f(\boldsymbol{x})}{\partial x_1^2} \quad |H_2| = \begin{vmatrix} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} \end{vmatrix} \quad |H_3| = \begin{vmatrix} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_3 \partial x_1} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_3 \partial x_2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_3^2} \end{vmatrix}$$

Teorema 2.7. (Condiciones necesarias de 2^{do} orden)
Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A un subconjunto abierto de $\mathbb{R}^n$ y f de clase $\mathcal{C}^2$. Si $\boldsymbol{x}_0$ es un mínimo local de f entonces la matriz $Hf(\boldsymbol{x}_0)$ es semidefinida positiva. Si $\boldsymbol{x}_0$ es un máximo local de f, entonces $Hf(\boldsymbol{x}_0)$ es semidefinida negativa.

Demostración. Sea $\boldsymbol{h} \in \mathbb{R}^n$ cualquiera, si $\boldsymbol{x}_0$ es un mínimo local entonces usando la condición necesaria de primer orden (teorema 2.4) y el teorema de Taylor de orden 2 (teorema 2.3) se tiene para todo $t \in \mathbb{R}$

$$f(\boldsymbol{x}_0 + t\boldsymbol{h}) - f(\boldsymbol{x}_0) = \frac{1}{2} t^2 \boldsymbol{h}^T H f(\boldsymbol{x}_0) \boldsymbol{h} + R_2(\boldsymbol{x}_0, t\boldsymbol{h})$$

Luego, por la propiedad que satisface el resto se tiene que

$$\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{h}) - f(\boldsymbol{x}_0)}{||t\boldsymbol{h}||^2} = \lim_{t\to 0} \frac{\frac{1}{2}t^2\boldsymbol{h}^T Hf(\boldsymbol{x}_0)\boldsymbol{h}}{||t\boldsymbol{h}||^2}$$

El límite de la izquierda es positivo pues $\boldsymbol{x}_0$ es mínimo local de f mientras que la expresión de la derecha no depende de t y es igual a

$$\frac{\frac{1}{2}\boldsymbol{h}^T Hf(\boldsymbol{x}_0)\boldsymbol{h}}{||\boldsymbol{h}||^2}$$

Esta cantidad es positiva sí y sólo sí $\boldsymbol{h}^T Hf(\boldsymbol{x}_0)\boldsymbol{h} \geq 0$. Como $\boldsymbol{h}$ es arbitrario, concluímos que $Hf(\boldsymbol{x}_0)$ es semidefinida positiva.

Para terminar la demostración notemos que $\boldsymbol{x}_0$ es máximo local de f sí y sólo sí $\boldsymbol{x}_0$ es un mínimo local de $-f$. Luego, la matriz Hessiana de $-f$, que es igual a $-Hf(\boldsymbol{x}_0)$, es semidefinida positiva, lo cual es equivalente a que $Hf(\boldsymbol{x}_0)$ sea semidefinida negativa. ■

Teorema 2.8. (Condiciones suficientes de 2^{do} orden)
Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ con A un subconjunto abierto de $\mathbb{R}^n$ y f de clase $\mathcal{C}^2$. Sea $\boldsymbol{x}_0 \in A$ un punto crítico de f.

1. *Si $Hf(\boldsymbol{x}_0)$ es definida positiva entonces $\boldsymbol{x}_0$ es un mínimo local de f.*
2. *Si $Hf(\boldsymbol{x}_0)$ es definida negativa entonces $\boldsymbol{x}_0$ es un máximo local de f.*

Demostración. Demostremos solamente la primera parte, la otra se demuestra de forma similar.

Si $\boldsymbol{x}_0$ es punto crítico, entonces usando el teorema de Taylor de orden 2 (teorema 2.3) se tiene que:

$$f(\boldsymbol{x}_0 + \boldsymbol{d}) - f(\boldsymbol{x}_0) = Hf(\boldsymbol{x}_0)(\boldsymbol{d}) + R_2(\boldsymbol{x}_0, \boldsymbol{d})$$

donde $R_2(\boldsymbol{x}_0, \boldsymbol{d}) \to 0$ cuando $\boldsymbol{d} \to 0$.

Como $Hf(\boldsymbol{x}_0)$ es definida positiva entonces existe $c > 0$ tal que

$$Hf(\boldsymbol{x}_0)(\boldsymbol{d}) \geq c\|\boldsymbol{d}\|^2 \ \forall \boldsymbol{d} \in \mathbb{R}^n$$

y como $\frac{R_2(\boldsymbol{x}_0,\boldsymbol{d})}{\|d\|^2} \to 0$ cuando $\boldsymbol{d} \to 0$, existe $\delta > 0$ tal que si $0 < \|\boldsymbol{d}\| < \delta$ entonces

$$|R_2(\boldsymbol{x}_0, \boldsymbol{d})| < c\|\boldsymbol{d}\|^2$$

y por tanto

$$f(\boldsymbol{x}_0 + \boldsymbol{d}) - f(\boldsymbol{x}_0) > c\|\boldsymbol{d}\|^2 - c\|\boldsymbol{d}\|^2 = 0$$

para todo $0 < \|\boldsymbol{d}\| < \delta$, lo que implica que $\boldsymbol{x}_0$ es un mínimo local. ■

Ejemplo 2.6. Encuentre los puntos críticos de la siguiente función y clasifíquelos.

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

Solución.

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{x^2+y^2+1}2x = 0 \quad \Rightarrow \quad x = 0 \\ \frac{\partial f}{\partial y} &= \frac{1}{x^2+y^2+1}2y = 0 \quad \Rightarrow \quad y = 0 \end{aligned}$$

y por tanto el único punto crítico es $(x,y)=(0,0)$. Veamos ahora la matriz Hessiana de la función f evaluada en este punto:

$$\left.\frac{\partial^2 f}{\partial x^2}\right|_{(0,0)} = \left.\frac{-x^2+y^2+2}{(x^2+y^2+1)^2}\right|_{(0,0)} = 2$$

$$\left.\frac{\partial^2 f}{\partial x\partial y}\right|_{(0,0)} = \left.\frac{-4xy}{(x^2+y^2+1)^2}\right|_{(0,0)} = 0$$

$$\left.\frac{\partial^2 f}{\partial y^2}\right|_{(0,0)} = \left.\frac{x^2-y^2+2}{(x^2+y^2+1)^2}\right|_{(0,0)} = 2$$

es decir, la matriz Hessiana queda:

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

la cual es trivialmente definida positiva y por tanto el punto $(0,0)$ es un mínimo local de f. □

Ejemplo 2.7. Dada la función $f(x,y) = y - 4x^2 + 3xy - y^2$, podemos saber si la función presenta un máximo o un mínimo a partir de la Matriz Hessiana.

Solución.

$$\frac{\partial f(x,y)}{\partial x} = -8x+3y \qquad \frac{\partial f(x,y)}{\partial y} = 1+3x-2y$$

$$\frac{\partial f(x,y)}{\partial x^2} = -8 \qquad \frac{\partial^2 f(x,y)}{\partial y^2} = -2$$

$$\frac{\partial^2 f(x,y)}{\partial y\partial x} = 3 \qquad \frac{\partial^2 f(x,y)}{\partial x\partial y} = 3$$

Luego,

$$H = \begin{pmatrix} -8 & 3 \\ 3 & -2 \end{pmatrix}$$

Se tiene que las submatrices corresponden a

$$|H_1| = -8 < 0 \qquad |H_2| = \begin{vmatrix} -8 & 3 \\ 3 & -2 \end{vmatrix} = 16 - 9 = 7 > 0$$

As, se tiene que la matriz Hessiana es constante para todos los puntos. Al ver los signos de los determinantes de las submatrices, se observa que estos son alternados y que $|H_1|$ es negativo. Luego, la matriz Hessiana es definida negativa y todo punto crítico de f será un máximo. □

2.3. Funciones convexas

Motivación: La familia de funciones convexas a valores reales corresponde a una familia mucho más amplia que la de funciones lineales. Lo que nos interesa es caracterizar esta familia, en especial el caso de las funciones convexas diferenciables, y presentar condiciones necesarias y suficientes para que un determinado elemento de un espacio vectorial normado sea el mínimo de la función convexa sobre algún conjunto. La caracterización que haremos de las funciones convexas es extensible a las funciones cóncavas.

Definición 2.6. Una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, con D convexo, es convexa si

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D, \quad \forall \lambda \in [0,1]$$

Definición 2.7. Una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$, con D convexo, es estrictamente convexa si

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) < \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D \quad \forall \lambda \in (0,1)$$

De manera más intuitiva, una función f es convexa si el segmento que une dos puntos pertenecientes al grafo de la función se ubica por sobre el grafo.
Determinar la convexidad de una función por medio de esta definición resulta complicado por lo que recurriremos a algunos criterios de diferenciabilidad cuando estos son aplicables.

Nota 2.5. Para que esta definición tenga sentido necesitamos que el dominio de la función sea un conjunto convexo. Esto es importante porque la definición que presentamos requiere que, para dos puntos cualesquiera $\boldsymbol{x}$ e $\boldsymbol{y}$ en el dominio de la función, las combinaciones convexas de $\boldsymbol{x}$ e $\boldsymbol{y}$, es decir los puntos de la forma $\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$, estén en el dominio de la función, lo que ocurre si el dominio es un conjunto convexo. A partir de ahora, cada vez que nos refiramos a una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ convexa entenderemos que D es igual a todo el espacio vectorial $\mathbb{R}^n$, que es convexo, o si es sólo una parte de $\mathbb{R}^n$, D será un subconjunto convexo de $\mathbb{R}^n$.

Definición 2.8. El hipografo de una función corresponde a todos los puntos situados bajo el grafo de la función y queda definido por el conjunto

$$E = \{(\boldsymbol{x}, c) : \boldsymbol{x} \in D, c \in \mathbb{R}, f(\boldsymbol{x}) \geq c\}$$

Definición 2.9. El epígrafo de una función corresponde a todos los puntos situados sobre el grafo de la función y queda definido por el conjunto

$$E = \{(\boldsymbol{x}, c) : \boldsymbol{x} \in D, c \in \mathbb{R}, f(\boldsymbol{x}) \leq c\}$$

El epígrafo de una función nos permite relacionar funciones convexas con conjuntos convexos.

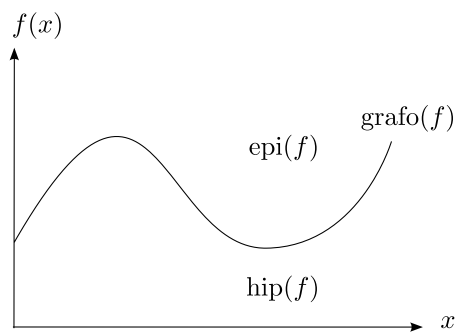


Figura 2.1: Epígrafo, hipografo y grafo de una función

A partir de la definición 2.9 podemos obtener otro criterio de convexidad.

Teorema 2.9. *$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ es convexa si y sólo si el epígrafo de f es convexo en $\mathbb{R}^{n+1}$. Esto es, $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $c_1, c_2 \in \mathbb{R}$ tales que $(\boldsymbol{x}, c_1), (\boldsymbol{y}, c_2) \in epi(f)$ se tiene que*

$$\lambda(\boldsymbol{x}, c_1) + (1-\lambda)(\boldsymbol{y}, c_2) \in epi(f)$$

Demostración.
($\Rightarrow$): Supongamos que $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función convexa, entonces debemos demostrar que $epi(f)$ es convexo. Sean $\boldsymbol{x}, \boldsymbol{y} \in D,\ c_1, c_2 \in \mathbb{R}$ tales que $(\boldsymbol{x}, c_1), (\boldsymbol{y}, c_2) \in epi(f)$, es decir:

$$f(\boldsymbol{x}) \leq c_1 \ , \ f(\boldsymbol{y}) \leq c_2 \tag{*}$$

Sea $\lambda \in [0, 1]$ cualquiera, por definición de convexidad tenemos que:

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Usando (*), se obtiene que

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda c_1 + (1-\lambda) c_2$$

Lo que es equivalente a

$$(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda c_1 + (1-\lambda)c_2) = \lambda(\boldsymbol{x}, c_1) + (1-\lambda)(\boldsymbol{y}, c_2) \in epi(f)$$

Concluímos que $epi(f)$ es convexo.

($\Leftarrow$): Supongamos que $epi(f)$ es convexo, entonces debemos demostrar que $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ es convexa. Sean $\boldsymbol{x}, \boldsymbol{y} \in D$ y notemos que $(\boldsymbol{x}, f(\boldsymbol{x})), (\boldsymbol{y}, f(\boldsymbol{y})) \in epi(f)$. Como $epi(f)$ es convexo, para todo $\lambda \in [0, 1]$ el punto $\lambda(\boldsymbol{x}, f(\boldsymbol{x})) + (1-\lambda)(\boldsymbol{y}, f(\boldsymbol{y}))$ pertenece a $epi(f)$. Esto último es equivalente a:

$$\begin{aligned} & (\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})) \in epi(f) \\ \iff \quad & f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \end{aligned}$$

Por lo tanto concluímos que f es convexa. ■

Teorema 2.10. *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función convexa. Entonces todo mínimo local de f en A es mínimo global de f en D.*

Demostración. Sea $\boldsymbol{x}_0$ un mínimo local de f en D y supongamos que $\boldsymbol{x}_0$ no es mínimo global de f, es decir, existe $\boldsymbol{x} \in D$ tal que $f(\boldsymbol{x}) < f(\boldsymbol{x}_0)$. Como D es convexo se tiene que

$$\forall \lambda \in (0, 1) \qquad (1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{x} \in D$$

Por lo tanto, en el segmento que une $\boldsymbol{x}$ y $\boldsymbol{x_0}$ se cumple:

$$\begin{aligned} f((1-\lambda)\boldsymbol{x_0} + \lambda\boldsymbol{x}) &\leq (1-\lambda) f(\boldsymbol{x}_0) + \lambda f(\boldsymbol{x}) \\ &< (1-\lambda) f(\boldsymbol{x}_0) + \lambda f(\boldsymbol{x}_0) \\ &= f(\boldsymbol{x}_0) \end{aligned}$$

Esto genera una contradicción pues en la medida que $\lambda \to 0$, se tendrá que $(1-\lambda)\boldsymbol{x}_0 + \lambda\boldsymbol{x} \to \boldsymbol{x}_0$. Luego, como $\boldsymbol{x}_0$ es mínimo local y f es convexa, para $\lambda \sim 0$ se debe cumplir que

$$f(\boldsymbol{x_0}) \leq f((1-\lambda)\boldsymbol{x_0} + \lambda\boldsymbol{x}).$$

■

Teorema 2.11. *Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable con D abierto y convexo. f es convexa (respectivamente estrictamente convexa) si y sólo si*

$$f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D \quad (resp.\ >) \tag{2.2}$$

Demostración.
($\Rightarrow$): Supongamos que f es convexa, entonces

$$\begin{aligned}
& f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) && \leq && \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) && \forall \boldsymbol{x}, \boldsymbol{y} \in D,\ \lambda \in [0,1] \\
\iff \quad & f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y})) && \leq && \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) && \forall \boldsymbol{x}, \boldsymbol{y} \in D,\ \lambda \in [0,1] \\
\iff \quad & \frac{f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y})) - f(\boldsymbol{y})}{\lambda} && \leq && f(\boldsymbol{x}) - f(\boldsymbol{y}) && \forall \boldsymbol{x}, \boldsymbol{y} \in D,\ \lambda \in]0,1]
\end{aligned}$$

Aplicando límite cuando $\lambda \to 0^+$

$$\begin{aligned}
& \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y}) \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) \\
\iff \quad & f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y})
\end{aligned}$$

($\Leftarrow$): Sean $\boldsymbol{x}, \boldsymbol{y} \in D$, $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ y $\lambda \in [0,1]$. Supongamos que f cumple (2.2), entonces

$$\begin{gathered}
\lambda f(\boldsymbol{x}) \geq \lambda (f(\boldsymbol{z}) + \nabla f(\boldsymbol{z}) \cdot (\boldsymbol{x}-\boldsymbol{z})) \\
(1-\lambda) f(\boldsymbol{y}) \geq (1-\lambda)(f(\boldsymbol{z}) + \nabla f(\boldsymbol{z}) \cdot (\boldsymbol{y}-\boldsymbol{z}))
\end{gathered}$$

Sumando las últimas dos desigualdades se obtiene que

$$\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \geq f(\boldsymbol{z}) + \nabla f(\boldsymbol{z}) \cdot (\underbrace{\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} - \boldsymbol{z}}_{0})$$

Es decir

$$\begin{aligned}
& f(z) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \\
\iff \quad & f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})
\end{aligned}$$

■

Teorema 2.12. *Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable con D abierto y convexo. f es convexa (respectivamente estrictamente convexa) si y sólo si*

$$(\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})) \cdot (\boldsymbol{x}-\boldsymbol{y}) \geq 0 \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D \quad (\textit{resp.}\ >) \tag{2.3}$$

Demostración.
($\Rightarrow$): Supongamos que f es convexa. De acuerdo al teorema 2.10 se cumplen

$$\begin{gathered}
f(\boldsymbol{x}) \geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y}) \\
f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x}) \cdot (\boldsymbol{y}-\boldsymbol{x})
\end{gathered}$$

Sumando estas desigualdades obtenemos

$$\begin{aligned}
& 0 \geq \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x}-\boldsymbol{y}) + \nabla f(\boldsymbol{x}) \cdot (\boldsymbol{y}-\boldsymbol{x}) \\
\iff \quad & (\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})) \cdot (\boldsymbol{x}-\boldsymbol{y}) \geq 0
\end{aligned}$$

($\Leftarrow$): Sean $\boldsymbol{x}, \boldsymbol{y} \in A$. Definamos $g : [0,1] \to \mathbb{R}$ como $g(\lambda) = f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y}))$. Como g es derivable obtenemos $g'(\lambda) = \nabla f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y})) \cdot (\boldsymbol{x}-\boldsymbol{y})$.
Supongamos que se cumple (2.3), entonces para todo $\lambda \in (0,1)$

$$\begin{aligned}
& g'(\lambda) - g'(0) = (\nabla f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y})) - \nabla f(\boldsymbol{y})) \cdot (\boldsymbol{x}-\boldsymbol{y}) \\
& g'(\lambda) - g'(0) = (\nabla f(\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y})) - \nabla f(\boldsymbol{y})) \cdot \left(\frac{\boldsymbol{y} + \lambda(\boldsymbol{x}-\boldsymbol{y}) - \boldsymbol{y}}{\lambda} \right) \\
& g'(\lambda) - g'(0) \geq 0
\end{aligned}$$

Aplicando el teorema del valor medio (teorema 1.23) a g, existe $\tilde{\lambda} \in (0,1)$ tal que

$$g(1) - g(0) = g'(\tilde{\lambda}) \cdot 1$$

Pero $g'(\tilde{\lambda}) \geq g'(0)$, luego

$$g(1) - g(0) \geq g'(0)$$

Reemplazando, se obtiene que

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{y}) &\geq \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \\ f(\boldsymbol{x}) &\geq f(\boldsymbol{y}) + \nabla f(\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \end{aligned}$$

Usando el teorema 2.9 concluímos que f es convexa. ■

Teorema 2.13. *Sea $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase $\mathcal{C}^2$ con D abierto y convexo. f es convexa (respectivamente estrictamente convexa) en D si y sólo si $Hf(\boldsymbol{x})$ es semi-definida positiva (respectivamente definida positiva) para todo $\boldsymbol{x} \in D$.*

Demostración.
($\Rightarrow$): Supongamos que f es convexa. Sean $\boldsymbol{x} \in D$, $\boldsymbol{h} \in \mathbb{R}^n \setminus \{0\}$ cualesquiera y definamos $g : [0, \varepsilon] \to \mathbb{R}$, para valores de λ tal que $\boldsymbol{x} + \lambda \boldsymbol{h} \in D$, como $g(\lambda) = \nabla f(\boldsymbol{x} + \lambda \boldsymbol{h}) \cdot \boldsymbol{h}$. Como g es derivable se obtiene

$$g'(\lambda) = \boldsymbol{h}^T Hf(\boldsymbol{x} + \lambda \boldsymbol{h})\boldsymbol{h}$$

debemos demostrar que

$$g'(0) = \lim_{\lambda \to 0^+} \frac{g(\lambda) - g(0)}{\lambda} \geq 0$$

Del teorema 2.11 tenemos que

$$\begin{aligned} g(\lambda) - g(0) &= (\nabla f(\boldsymbol{x} + \lambda \boldsymbol{h}) - \nabla f(\boldsymbol{x})) \cdot \boldsymbol{h} \\ &= (\nabla f(\boldsymbol{x} + \lambda \boldsymbol{h}) - \nabla f(\boldsymbol{x})) \cdot \left(\frac{\boldsymbol{x} + \lambda \boldsymbol{h} - \boldsymbol{x}}{\lambda} \right) \\ &\geq 0 \end{aligned}$$

Por lo tanto se tiene que $g'(0) = \boldsymbol{h}^T Hf(\boldsymbol{x})\boldsymbol{h} \geq 0$. Como $\boldsymbol{h}$ es arbitrario, se concluye que $Hf(\boldsymbol{x})$ es semi-definida positiva.
($\Leftarrow$): Sean $\boldsymbol{x}, \boldsymbol{y} \in D$ cualesqueria, por el teorema 2.3 se tiene que

$$f(\boldsymbol{x}) = f(\boldsymbol{y}) + \nabla f(\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{y}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{y})^T Hf(\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{y}) + R_2(\boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y})$$

Luego, si $Hf(\boldsymbol{y})$ es semi-definida positiva se obtiene

$$f(\boldsymbol{x}) - f(\boldsymbol{y}) - \nabla f(\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{y}) - R_2(\boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}) \geq 0$$

de acuerdo al teorema 2.10 esta última desigualdad verifica que f es convexa. ■

La convexidad local es una condición necesaria para la existencia de mínimos locales de funciones y la convexidad estricta es una condición suficiente para la existencia de un mínimo global. Para una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ diferenciable el teorema 2.13 nos asegura provee un criterio útil para determinar la convexidad de una función cuando esta es diferenciable.

Ejemplo 2.8. $-\ln(x)$ con $x \in \mathbb{R}_{++}$ $(x > 0)$ es convexa. En efecto, mediante la diferenciabilidad de la función podemos garantizar que es convexa ya que $\frac{\partial^2 f}{\partial x^2} = \frac{1}{x^2} > 0$ $\forall x > 0$.

Ejemplo 2.9. Las funciones en $\mathbb{R}$:

1. $ax + b$ de $\mathbb{R}$ en $\mathbb{R}$ con cualquier $a, b \in \mathbb{R}$
2. $\exp(\alpha x)$ de $\mathbb{R}$ en $\mathbb{R}$ con cualquier $a \in \mathbb{R}$
3. x^α de $\mathbb{R}_+$ en $\mathbb{R}_+$ con $\alpha \geq 1$
4. $|x|^\alpha$ de $\mathbb{R}$ en $\mathbb{R}_+$ con $\alpha \geq 1$

Las funciones en $\mathbb{R}^2$:

1. $a^T x + b$ de $\mathbb{R}^2$ en $\mathbb{R}$ con $a \in \mathbb{R}^2$, $b \in \mathbb{R}$
2. $A(\alpha x_1 + \beta x_2)^\rho$ de $\mathbb{R}^2_+$ en $\mathbb{R}$ con $A, \alpha, \beta > 0$ y $\rho > 1$
3. $x_1^\alpha x_2^\beta$ de $\mathbb{R}^2_+$ en $\mathbb{R}$ con $\alpha + \beta > 1$

Las normas en $\mathbb{R}^n$ y toda función $f(x) = a^T x + b$ con $a, x \in \mathbb{R}^n$ y $b \in \mathbb{R}$ son funciones convexas.

2.4. Funciones cóncavas

Definición 2.10. Una funcion $f : D \subset \mathbb{R}^n \to \mathbb{R}$, con D abierto y convexo, es cóncava (respectivamente estrictamente cóncava) si la función $-f$ es convexa (respectivamente estrictamente convexa). De esta forma, f es cóncava (resp. estrictamente cóncava) si

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in D, \quad \forall \lambda \in (0,1)$$

Alternativamente, una función es cóncava si su hipografo es convexo, esto es $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ c_1, c_2 \in \mathbb{R}$ tales que $(\boldsymbol{x}, c_1), (\boldsymbol{y}, c_2) \in hip(f)$ se tiene que

$$\lambda(\boldsymbol{x}, f(\boldsymbol{x})) + (1-\lambda)(\boldsymbol{y}, f(\boldsymbol{y})) \in hip(f)$$

Nota 2.6. Seguimos suponiendo que el dominio de la función es convexo por las mismas razones consideradas para las funciones convexas. Son tres hechos completamente distintos pero relacionados: la convexidad del dominio, la cóncavidad de la función y la convexidad del hipografo en el caso de las funciones cóncavas.

De la definición presentada tenemos que todo lo expuesto sobre funciones convexas es válido cambiando f por $-f$. Así, en los teoremas 2.9, 2.10, 2.11, 2.12 y 2.13 nos basta con reemplazar mínimo por máximo, invertir las desigualdades, cambiar convexo por cóncavo, (semi) definida positiva por (semi) definida negativa y cambiar epígrafo por hipografo para obtener el caso cóncavo.

La concavidad es una condición necesaria para la existencia de máximos locales de funciones y la concavidad estricta es una condición suficiente para la existencia un máximo global. Para una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ el teorema 2.13 nos provee un criterio útil para determinar la concavidad de una función cuando esta es diferenciable.

2.5. Extremos restringidos y Multiplicadores de Lagrange

Motivación: Veremos intuitivamente el caso en que aparecen restricciones en un problema de maximización o minimización. Retomaremos esto más adelante pero es conveniente tener al menos la noción de como afrontar estos problemas. Las demostraciones detalladas las veremos en el capítulo final para seguir con temas tanto o más importantes.

Teorema 2.14. *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ una función continua y A un conjunto cerrado y acotado. Entonces f alcanza su mínimo y su máximo en A, es decir, existen $\boldsymbol{x_0}, \boldsymbol{x_1} \in A$ tales que:*

$$f(\boldsymbol{x_0}) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{x_1}) \quad \forall \boldsymbol{x} \in A$$

Demostración. Para el caso del mínimo, sea

$$m = \inf_{\boldsymbol{x} \in A} f(\boldsymbol{x})$$

por definición del ínfimo, sabemos que existe una sucesión $\{\boldsymbol{x_k}\} \subset A$ tal que $f(\boldsymbol{x_k}) \xrightarrow{k\to\infty} m$. Demostremos que esta sucesión, por estar contenida en A, tiene una subsucesión convergente. En efecto, como A es acotado, existe $K > 0$ tal que $||\boldsymbol{x}|| \leq K \quad \forall \boldsymbol{x} \in A$, luego, para todo $k \in \mathbb{N}$ se cumple que $||\boldsymbol{x_k}|| \leq K$. Tomemos ahora la sucesión real de las primeras coordenadas, es decir para todo $k \in \mathbb{N}$ definimos $\{y_k^1\}_{k\in\mathbb{N}}$ por

$$y_k^1 = \boldsymbol{x_k^1}$$

Notemos que $|y_k^1| \leq ||\boldsymbol{x_k}|| \leq K$, por lo tanto y_k^1 es acotada y posee una subsucesión convergente $y_{k_j}^1 \xrightarrow{j\to\infty} y^1$. A continuación tomemos la sucesión de las segundas coordenadas de la subsucesión definida por los índices anteriores, es decir para todo $j \in \mathbb{N}$ definimos $\{y_j^2\}_{j\in\mathbb{N}}$ por

$$y_j^2 = \boldsymbol{x_{k_j}^2}$$

Notemos que $|y_j^2| \leq ||\boldsymbol{x_{k_j}}|| \leq K$, por lo que y_j^2 es acotada y posee una subsucesión convergente $y_{j_l}^2 \xrightarrow{l\to\infty} y^2$, además, como los índices k_{j_l} están contenidos en los índices k_j se tiene que $y_{k_{j_l}}^1 \xrightarrow{l\to\infty} y^1$. Es decir,

$$\boldsymbol{x_{k_{j_l}}^1} \xrightarrow{l\to\infty} y^1$$

$$\boldsymbol{x_{k_{j_l}}^2} \xrightarrow{l\to\infty} y^2$$

Repitiendo este proceso hasta la última coordenada, obtendremos que existe una subsucesión tal que cada coordenada de $\{\boldsymbol{x_k}\}$ es convergente, es decir

$$\boldsymbol{x_{k_i}^1} \xrightarrow{i\to\infty} y^1$$

$$\boldsymbol{x_{k_i}^2} \xrightarrow{i\to\infty} y^2$$

$$\vdots$$

$$\boldsymbol{x_{k_i}^n} \xrightarrow{i\to\infty} y^n$$

O equivalentemente

$$\boldsymbol{x_{k_i}} \xrightarrow{i\to\infty} \boldsymbol{y} = (y^1, \ldots, y^n)^T$$

Como A es cerrado, necesariamente $\boldsymbol{y} \in A$ y como f es continua, se tiene que

$$\lim_{i\to\infty} f(\boldsymbol{x_{k_i}}) = f(\boldsymbol{y}) = m$$

Por lo tanto $m > -\infty$ y en el punto $\boldsymbol{y}$ se alcanza el ínfimo de f en A. Luego,

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in A$$

La demostración para el caso del máximo es análoga. ∎

2.5.1. Condiciones de 1er orden para extremos restringidos

Teorema 2.15. (Teorema de los multiplicadores de Lagrange, forma geométrica)
Sean $f : A \subset \mathbb{R}^n \to \mathbb{R}$ y $g : A \subset \mathbb{R}^n \to \mathbb{R}$ funciones diferenciables. Sea $\boldsymbol{x}_0 \in A$ tal que $g(\boldsymbol{x}_0) = c$, y sea S el conjunto de nivel para g con valor c, es decir

$$S = \{\boldsymbol{x} \in A : g(\boldsymbol{x}) = c\}$$

Supongamos que $\nabla g(\boldsymbol{x}_0) \neq 0$ y que $f|_S$ (f restringida a S) tiene un mínimo (o máximo) local en S en $\boldsymbol{x}_0$, es decir, $\boldsymbol{x}_0$ es una solución del problema:

$$\begin{array}{lll} \min\limits_{x} & f(\boldsymbol{x}) \\ s.a & g(\boldsymbol{x}) = c \end{array} \qquad \text{ó} \qquad \begin{array}{ll} \max\limits_{x} & f(\boldsymbol{x}) \\ s.a & g(\boldsymbol{x}) = c \end{array}$$

entonces existe un valor $\lambda \in \mathbb{R}$, llamado multiplicador de Lagrange, tal que $(\boldsymbol{x}_0, \lambda)$ es mínimo (o máximo) local de la función Lagrangeano:

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda(g(\boldsymbol{x}) - c)$$

sobre la cual podemos aplicar la condición de primer orden y obtener que λ satisface

$$\nabla f(\boldsymbol{x}_0) = \lambda \nabla g(\boldsymbol{x}_0)$$

Demostración. Supongamos que $\boldsymbol{x}_0$ es solución del problema de minimización (el otro caso es similar) y sea

$$S = \{\boldsymbol{x} \in A : g(\boldsymbol{x}) = c\}$$

Como $\nabla g(\boldsymbol{x}_0) \neq \boldsymbol{0}$, entonces este vector es normal a la superficie S, por otro lado, el plano tangente a S en $\boldsymbol{x}_0$ se caracteriza como

$$\pi : \ \nabla g(\boldsymbol{x}_0)^T \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0$$

o equivalentemente

$$\pi = \{\sigma'(0) : \sigma : \mathbb{R} \to A, \ \sigma(t) \in S \ \forall t, \ \sigma(0) = \boldsymbol{x}_0\}$$

Entonces, siendo $\boldsymbol{x}_0$ un mínimo de f, la función $t \longmapsto f(\sigma(t))$ tiene un mínimo en 0 para cada σ, por tanto

$$\nabla f(\boldsymbol{x}_0)^T \cdot \sigma'(0) = 0$$

es decir, $\nabla f(\boldsymbol{x}_0)$ es ortogonal al plano tangente o lo que es lo mismo, paralelo al vector $\nabla g(\boldsymbol{x}_0)$. Por lo tanto, existe un real λ tal que

$$\nabla f(\boldsymbol{x}_0) = \lambda \nabla g(\boldsymbol{x}_0)$$

■

Observemos que la función Lagrangeano consta de $n+1$ variables (el multiplicador es una variable más) y que el teorema provee una condición necesaria que es equivalente a que el siguiente sistema tenga solución:

$$\begin{cases} \nabla f(\boldsymbol{x}_0) = \lambda \nabla g(\boldsymbol{x}_0) \\ g(x_1, ..., x_n) = c \end{cases} \iff \begin{cases} \dfrac{\partial f}{\partial x_1}(x_1, ..., x_n) = \lambda \dfrac{\partial g}{\partial x_1}(x_1, ..., x_n) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(x_1, ..., x_n) = \lambda \dfrac{\partial g}{\partial x_n}(x_1, ..., x_n) \\ \\ g(x_1, ..., x_n) = c \end{cases} \tag{2.4}$$

Corolario 2.3. *Si f al restringirse a una superficie S, tiene un máximo o un mínimo en $\boldsymbol{x}_0$, entonces $\nabla f(\boldsymbol{x}_0)$ es perpendicular a S en $\boldsymbol{x}_0$.*

Ejemplo 2.10. Sea $S \subset \mathbb{R}^2$ la recta que pasa por el punto $(-1, 0)$ y tiene una inclinación de $45°$ y sea $f : \mathbb{R}^2 \to \mathbb{R}$ tal que $(x, y) \mapsto x^2 + y^2$. Hallar los puntos extremos de f sobre la recta S.

Solución. Veamos que S se puede escribir como

$$S = \{(x, y) : y - x - 1 = 0\}$$

y denotemos por (x_0, y_0) el posible candidato a ser punto extremo. En este caso $g(x, y) = y - x - 1$, $c = 0$. De esta forma, el Lagrangeano del problema corresponde a

$$L(x, y, \lambda) = x^2 + y^2 - \lambda(y - x - 1)$$

a partir del cual, aplicando el sistema Lagrangeano (2.4) se obtiene

$$\begin{cases} 2x_0 = -\lambda \\ 2y_0 = \lambda \\ y_0 = x_0 + 1 \end{cases} \Rightarrow \quad \begin{matrix} x_0 = -y_0 \\ y_0 = x_0 + 1 \end{matrix} \quad \Rightarrow \quad \begin{matrix} x_0 = -1/2 \\ y_0 = 1/2 \end{matrix}$$

Es decir, el extremo de f es $(x_0, y_0) = (-\frac{1}{2}, \frac{1}{2})$ y el valor del multiplicador de lagrange es $\lambda = 1$. □

El siguiente contraejemplo resulta interesante:

Ejemplo 2.11. Consideremos el problema

$$\begin{array}{cc} \min\limits_{x_1, x_2} & x_1 + x_2 \\ \text{s.a} & 2x_1 - 3x_2 = 0 \end{array}$$

Observemos que $\nabla f(x_1, x_2) = (1, 1)$ y $\nabla g(x_1, x_2) = (2, -3)$. Entonces, no existe $\lambda \in \mathbb{R}$ tal que el sistema que define las condiciones de primer orden del Lagrangeano tenga solución. Esto ocurre pues el problema original no tiene solución.

En general, para poder determinar si los puntos extremos son mínimos locales, máximos o puntos sillas debemos analizar el comportamiento de la 2^{da} derivada o usar otros argumentos, como veremos más adelante.

Supongamos ahora que tenemos k restricciones de igualdad, es decir

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : g_1(x_1, ..., x_n) = 0, \ldots, g_k(x_1, ..., x_n) = 0\}$$

y el problema de optimización es:

$$\begin{array}{lcl} \min\limits_{x} f(\boldsymbol{x}) & \quad ó \quad & \max\limits_{x} f(\boldsymbol{x}) \\ \boldsymbol{x} \in S & & \boldsymbol{x} \in S \end{array} \tag{2.5}$$

donde $f, g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in \{1, \dots, k\}$ son funciones diferenciables.

Entonces el teorema de los multiplicadores de Lagrange se extiende de la siguiente forma:

Teorema 2.16. *Si los problemas (2.5) tienen un mínimo o máximo local $\boldsymbol{x}_0$, es decir, existe una vecindad V de $\boldsymbol{x}_0$ tal que:*

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) \ \forall \boldsymbol{x} \in V \cap S \qquad (ó\ f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0))$$

entonces existen k números reales (multiplicadores) $\lambda_1, \dots, \lambda_k$ tales que:

$$\nabla f(\boldsymbol{x}_0) = \lambda_1 \nabla g_1(\boldsymbol{x}_0) + \ldots + \lambda_k \nabla g_k(\boldsymbol{x}_0)$$

siempre que los vectores $\nabla g_1(\boldsymbol{x}_0), \dots, \nabla g_k(\boldsymbol{x}_0)$ sean linealmente independientes.

Ejemplo 2.12. Encontrar el máximo valor de $f(x, y, z) = x + 2y + 3z$ en la curva de la intersección del plano $x - y + z = 1$ y la circunferencia $x^2 + y^2 = 1$.

Solución. Sabemos que el problema tiene solución pues se quiere maximizar una función continua sobre un conjunto compacto. Para encontrarla debemos plantear el Lagrangeano del problema, el cual corresponde a

$$L(x, y, z, \lambda_1, \lambda_2) = x + 2y + 3z - \lambda_1(x - y + z - 1) - \lambda_2(x^2 + y^2 - 1)$$

Aplicando el sistema lagrangeano (2.4) obtenemos

$$\left\{ \begin{array}{l} 1 = \lambda_1 + 2\lambda_2 x_0 \\ 2 = -\lambda_1 + 2\lambda_2 y_0 \\ 3 = \lambda_1 \\ x_0 - y_0 + z_0 = 1 \\ x_0^2 + y_0^2 = 1 \end{array} \right. \quad \Rightarrow \quad \begin{array}{l} 5x_0 = -2y_0 \\ x_0 - y_0 + z_0 = 1 \\ x_0^2 + y_0^2 = 1 \end{array} \quad \Rightarrow \quad \begin{array}{l} x_0 = \pm 2/\sqrt{29} \\ y_0 = \mp 5/\sqrt{29} \\ z_0 = 1 \mp 7/\sqrt{29} \end{array}$$

Es decir, f presenta dos extremos que son

$$(x_1, y_1, z_1) = \left(\frac{2}{29}, \frac{-5}{29}, 1 - \frac{7}{29} \right)$$

$$(x_2, y_2, z_2) = \left(\frac{-2}{29}, \frac{5}{29}, 1 + \frac{7}{29} \right)$$

mientras que los valores de los multiplicadores de Lagrange son $\lambda_1 = 3$ y $\lambda_2 = \mp\sqrt{29}/2$. Evaluando la función objetivo en estos dos puntos notamos que el primero corresponde al mínimo y el segundo al máximo. □

2.5.2. Condiciones de 2^{do} orden para extremos restringidos

Consideremos el siguiente problema:

$$\begin{array}{lllll} \min\limits_{x} & f(x_1, \dots, x_n) & \quad ó \quad & \max\limits_{x} & f(x_1, \dots, x_n) \\ \text{s.a} & g_i(x_1, \dots, x_n) = c & & \text{s.a} & g_i(x_1, \dots, x_n) = c \quad \forall i \in \{1, \dots, k\} \end{array} \tag{2.6}$$

donde f, g_i son funciones diferenciables.

Nota 2.7. En lo que sigue denotaremos $I = \{1, \ldots, k\}$.

Por el teorema anterior sabemos que si $\{\nabla g_i(x_1, ..., x_n)\}$ son l.i. entonces existe $\lambda \in \mathbb{R}^k$ tal que:

$$\nabla f(x_1, ..., x_n) = \lambda^T \nabla g(x_1, ..., x_n)$$

donde $g(x_1, ..., x_n) = (g_1(x_1, ..., x_n), ..., g_k(x_1, ..., x_n))^T$, es decir

$$\nabla L(x_1, ..., x_n, \lambda) = 0$$

donde L representa el Lagrangeano del problema (2.6).

De los puntos críticos para el Lagrangeano, más las restricciones, obtenemos posibles candidatos a mínimos o máximos locales:

$$\begin{aligned} \nabla L(\boldsymbol{x}) &= 0 \\ g_i(\boldsymbol{x}) &= 0 \ \forall i \in I \end{aligned}$$

En el caso sin restricciones, el criterio para determinar de que tipo eran los posibles extremos era analizando la matriz Hessiana de la función objetivo, en dependencia de si esta era definida negativa o positiva, entonces se tenía un máximo o un mínimo respectivamente.

Un criterio similar se tiene para el caso de un problema con restricciones, sin embargo no será necesario que el Hessiano, en este caso el del Lagrangeano, sea definido positivo o negativo para cada dirección $\boldsymbol{d}$, en realidad bastará que lo sea en un cierto conjunto que denominaremos conjunto de direcciones críticas y lo definiremos como:

$$K(\boldsymbol{x}) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I, \ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \leq 0\}$$

cuando el problema es de minimización y

$$K(\boldsymbol{x}) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I, \ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0\}$$

cuando es de maximización.

Teorema 2.17. *Sea $\boldsymbol{x}_0 \in S = \{\boldsymbol{x} : g_i(\boldsymbol{x}) = 0 \ \forall i \in I\}$. Supongamos que*

$$\{\nabla g_1(\boldsymbol{x}_0), \ldots, \nabla g_k(\boldsymbol{x}_0)\}$$

es linealmente independiente, entonces existe $\lambda \in \mathbb{R}^k$ tal que

$$\nabla L(\boldsymbol{x}_0, \lambda) = 0$$

Si además se tiene que

$$\boldsymbol{d}^T H_x L(\boldsymbol{x}_0, \lambda) \boldsymbol{d} > 0 \ \forall \boldsymbol{d} \in K(\boldsymbol{x}), \ \boldsymbol{d} \neq 0$$

entonces $\boldsymbol{x}_0$ es un mínimo local de (2.6). Mientras que si se tiene

$$\boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \lambda)) \boldsymbol{d} < 0 \ \forall \boldsymbol{d} \in K(\boldsymbol{x}), \ \boldsymbol{d} \neq 0$$

entonces $\boldsymbol{x}_0$ es un máximo local de (2.6).

Por razones pedagógicas, la demostración de este teorema no se hará hasta el capítulo 5 a fin de seguir con contenidos que son tanto o más importantes en el curso y hacerla cuando hayamos visto con detalle el teorema de los multiplicadores de Lagrange.

Hay que tener presente que en el teorema el hessiano del Lagrangeano es solo con respecto a $\boldsymbol{x}$, es decir

$$H_x L(\boldsymbol{x}_0, \lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial x_1^2}(\boldsymbol{x}_0, \lambda) & \frac{\partial^2 L}{\partial x_1 \partial x_2}(\boldsymbol{x}_0, \lambda) & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n}(\boldsymbol{x}_0, \lambda) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(\boldsymbol{x}_0, \lambda) & \frac{\partial^2 L}{\partial x_n \partial x_2}(\boldsymbol{x}_0, \lambda) & \cdots & \frac{\partial^2 L}{\partial x_n^2}(\boldsymbol{x}_0, \lambda) \end{bmatrix} = \left[\frac{\partial^2 L}{\partial x_i \partial x_j}(\boldsymbol{x}_0, \lambda) \right]_{i,j=1}^n$$

Ejemplo 2.13. En el ejemplo (2.10), diga si el punto $(x_0, y_0) = (-\frac{1}{2}, \frac{1}{2})$ es un máximo o un mínimo.

Solución. Para esto, construyamos primeramente el Lagrangeano del problema:

$$L(x, y, \lambda) = x^2 + y^2 - \lambda(y - x - 1)$$

Recordemos que el multiplicador de lagrange era $\lambda = 1$, por tanto, el Hessiano del Lagrangeano queda:

$$H_x L(x_0, y_0, \lambda) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

el cual es definido positivo para todo d, en particular para las direcciones del conjunto de direcciones críticas y por tanto el punto $(x_0, y_0) = (-\frac{1}{2}, \frac{1}{2})$ es un mínimo. □

Ejemplo 2.14. Optimice el valor de la función $f(x, y) = x$ sobre el conjunto de puntos (x, y) tales que $x^2 + 2y^2 = 3$.

Solución. El Lagrangeano para este problema queda de la siguiente forma:

$$L(x, y, \lambda) = x - \lambda(x^2 + 2y^2 - 3)$$

a partir del cual, aplicando el sistema lagrangeano (2.4) se obtiene

$$\begin{cases} 1 = 2\lambda x_0 \\ 4 = \lambda y_0 \\ x_0^2 + 2y_0^2 = 3 \end{cases}$$

$$\begin{cases} 1 = 2\lambda x_0 \\ 0 = 4\lambda y_0 \\ x_0^2 + 2y_0^2 = 3 \end{cases} \quad \Rightarrow \quad \begin{matrix} 1 = 2\lambda x_0 \\ x_0^2 + 2y_0^2 = 3 \end{matrix}$$

y por tanto, los posibles candidatos a extremos (λ no puede ser cero) son:

$$(x_1, y_1, \lambda_1) = \left(\sqrt{3}, 0, \frac{1}{2\sqrt{3}} \right)$$

$$(x_2, y_2, \lambda_2) = \left(-\sqrt{3}, 0, -\frac{1}{2\sqrt{3}} \right)$$

por otro lado se tiene que:

$$H_x L(x, y, \lambda) = \begin{pmatrix} -2\lambda & 0 \\ 0 & -4\lambda \end{pmatrix}$$

es decir, para

$$(x_1, y_1, \lambda_1) = \left(\sqrt{3}, 0, \frac{1}{2\sqrt{3}}\right)$$

$H_xL(x_1, y_1, \lambda_1)$ es definida negativa y por tanto $(x_1, y_1) = (\sqrt{3}, 0)$ es un máximo local mientras que para

$$(x_2, y_2, \lambda_2) = \left(-\sqrt{3}, 0, -\frac{1}{2\sqrt{3}}\right)$$

$H_xL(x_2, y_2, \lambda_2)$ es definida positiva y por lo tanto $(x_2, y_2) = (-\sqrt{3}, 0)$ es un mínimo local. $\square$

2.6. Ejercicios

Derivadas superiores y teorema de Taylor

Ejercicio 1. Sea $f(u,v,w)$ una función con derivadas parciales continuas de orden 1 y 2, y sea $g(x,y) = f(x+y, x-y, xy)$. Calcule $g_{xx} + g_{yy}$ en términos de derivadas de $f(u,v,w)$.

Ejercicio 2. Considere la funcíon $f(x,y) = x^3y + sen(x^2y)$ y verifique que

$$\frac{\partial^4 f}{\partial x \partial^2 y \partial x} = \frac{\partial^4 f}{\partial^2 y \partial^2 x}$$

Ejercicio 3. Encuentre una función $f : \mathbb{R}^2 \to \mathbb{R}$ continua tal que $\lim\limits_{\|\boldsymbol{x}\| \to +\infty} f(\boldsymbol{x}) = 0$ y que no posea máximo global.

Ejercicio 4. Sea $u(x,y)$ una función con derivadas parciales continuas de orden 2 y considere la función $v(s,t) = u(e^s \cos(t), e^s \operatorname{sen}(t))$. Demuestre que

$$\frac{\partial^2 v}{\partial s^2} + \frac{\partial^2 v}{\partial t^2} = e^{2s}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right)$$

Ejercicio 5. Considere la funcion $u : \mathbb{R}^{n+1} \to \mathbb{R}$, definida como

$$u(t,\boldsymbol{x}) = (\beta t)^\gamma \exp\left(\frac{-\alpha \|\boldsymbol{x}\|^2}{t}\right)$$

Muestre que u verifica la ecuacion de difusion:

$$\frac{\partial u}{\partial t}(t,\boldsymbol{x}) = k^2 \Delta u(t,\boldsymbol{x})$$

Para $k > 0$ y el laplaciano calculado sin considerar la derivada con respecto a t, si y sólo si $\alpha = \frac{1}{4k^2}$, $\gamma = -\frac{n}{2}$ y β cualquiera.

Ejercicio 6. Una función $f : \mathbb{R}^n \to \mathbb{R}$ se dice homogénea de grado m si:

$$f(tx) = t^m f(x) \quad \forall t \in \mathbb{R} \quad \forall x \in \mathbb{R}^n$$

1. Demuestre que si f es de clase $\mathcal{C}^1$ y homogénea de grado m entonces las derivadas parciales de f son homogéneas de grado $m-1$.

2. Demuestre que si f es de clase $\mathcal{C}^1$ entonces para todo $x \in \mathbb{R}^n$ se cumple:

$$mf(x) = \nabla f(x).$$

3. Demuestre que si f es de clase $\mathcal{C}^2$ entonces para todo $x \in \mathbb{R}^n$ se cumple:

$$m(m-1)f(x) = x^t Hf(x)x.$$

Ejercicio 7. Demostrar que si f es de clase $\mathcal{C}^2$ entonces

$$f(\boldsymbol{x}_0 + \boldsymbol{h}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\boldsymbol{h} + 1/2\boldsymbol{h}Hf(\boldsymbol{x}_0)\boldsymbol{h}^T + R_2(\boldsymbol{x}_0, \boldsymbol{h})$$

donde

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{R_2(\boldsymbol{x}_0, \boldsymbol{h})}{\|\boldsymbol{h}\|} = 0$$

Ejercicio 8. Una función $u = f(x, y)$ con segundas derivadas parciales continuas que satisfaga la ecuación de Laplace

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

se llama función armónica. Determinar cuales de las siguientes funciones son armónicas.

1. $u(x, y) = x^3 - 3xy^2$
2. $u(x, y) = \operatorname{sen}(x)\cosh(y)$
3. $u(x, y) = e^x seny$

Ejercicio 9. Sea $f : \mathbb{R}^2 \to \mathbb{R}$. Para cada x defina $g_x : \mathbb{R} \to \mathbb{R}$, $g_x(y) = f(x, y)$. Suponga que para cada x existe un único y tal que $g_x'(y) = 0$. Si se denota por $c(x)$ a tal punto y y se supone que la función c es diferenciable demostrar:

1. Si $\frac{\partial^2 f}{\partial y^2}(x, y) \neq 0$ para todo (x, y) entonces:

$$c'(x) = -\frac{\frac{\partial^2 f}{\partial y \partial x}(x, c(x))}{\frac{\partial^2 f}{\partial y^2}(x, c(x))}$$

2. Si $c'(x) = 0$, entonces existe un $\bar{y}$ tal que

$$\frac{\partial^2 f}{\partial y \partial x}(x, \bar{y}) = 0 \text{ y } \frac{\partial f}{\partial y}(x, \bar{y}) = 0$$

Ejercicio 10. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

$$f(x, y) = \begin{cases} x^2 \arctan(y/x) - y^2 \arctan(x/y) & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

1. Calcule $f(x, 0)$ y $f(0, y)$ para $x \neq 0$,$y \neq 0$ respectivamente.
2. Para $(x, y) \neq (0, 0)$, determine $\nabla f(x, y)$ y $H_f(x, y)$. ¿Es $H_f(x, y)$ matriz simétrica?.
3. ¿Se cumple que $\dfrac{\partial^2 f}{\partial x \partial y}(0, 0) = \dfrac{\partial^2 f}{\partial y \partial x}(0, 0)$? Justifique.
4. Encuentre el polinomio de Taylor de orden 2 en torno a (0,0).

Ejercicio 11. Calcular el polinomio de Taylor de orden 2 en torno a $(\frac{\pi}{2}, 1)$ de la función definida por

$$f(x, y) = \operatorname{sen}(xy) + \cos(xy) + 2(x + y)$$

Ejercicio 12. Calcular la expansión de Taylor de segundo orden de las funciones siguientes en los puntos señalados, y calcule una vecindad en torno al punto tal que la aproximación tenga un error de a lo más 10^{-2}.

1. $f(x,y,z) = (x^2+2xy+y^2)e^z$ en $\{(1,2,0),(3,2,5)\}$.
2. $f(x,y,z) = (x^3+3x^2y+y^3)e^{-z^2}$ en $\{(0,0,0),(3,2,3)\}$.
3. $f(x_1,x_2,x_3,x_4) = \log(\cos(x_1+x_2-x_3-x_4))$ en $(0,0,0,0)$.

Ejercicio 13. Encuentre la aproximación de primer orden $P_1(x,y,z)$ y la aproximación de segundo orden $P_2(x,y,z)$ para la función

$$f(x,y,z) = xe^y + ze^{2y}$$

en torno al punto $(x_0,y_0,z_0) = (0,0,0)$. Encuentre una vecindad en torno al origen que garantice un error de a lo más 10^{-3}.

Ejercicio 14. Sea $f:\mathbb{R}^n \to \mathbb{R}$ de clase $\mathcal{C}^\infty$. Pruebe que en general no se tiene que la serie de Taylor de f converge a f. Para esto considere como contraejemplo la función $f:\mathbb{R}\to\mathbb{R}$ definida por

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{si } x < 0 \\ 0 & \text{si } x \geq 0 \end{cases}$$

Pruebe que es $\mathcal{C}^\infty$ y estudie su serie de Taylor en torno a cero.

Ejercicio 15. Escribir la fórmula general de Taylor de orden 3.

Extremos de funciones con valores reales

Ejercicio 16. Sea $f:\mathbb{R}^n \to \mathbb{R}$ una función continua. Pruebe que sobre todo conjunto compacto f alcanza un mínimo y un máximo.

Ejercicio 17. Sea $f:\mathbb{R}^n \to \mathbb{R}$ una función continua que satisface $\lim\limits_{\|\boldsymbol{x}\|\to+\infty} f(\boldsymbol{x}) = \infty$. Pruebe que f posee un mínimo global.

Ejercicio 18. Sea $f:\mathbb{R}^n \to \mathbb{R}$ una función continua que satisface $\lim\limits_{\|\boldsymbol{x}\|\to+\infty} f(\boldsymbol{x}) = 0$ y $f(\boldsymbol{x}_0) > 0$ para algún $\boldsymbol{x}_0 \in \mathbb{R}^n$. Pruebe que f posee un máximo global.

Ejercicio 19. Sea $f(x,y) = (ax^2+by^2)e^{-(x^2+y^2)}$, con $a,b > 0$. Calcule todos los máximos y mínimos globales de f.

Indicación. Analice los casos $a=b$,$a<b$ y $a>b$.

Ejercicio 20. Calcular los extremos relativos y los puntos silla para las funciones

1. $\cos(x)\cosh(y)$ donde $\cosh(x) = \frac{1}{2}(e^x+e^{-x})$.
2. $f(\boldsymbol{x}) = \sum_{i=0}^m \|\boldsymbol{x}-\boldsymbol{y}^i\|^2$ donde $\boldsymbol{x}\in\mathbb{R}^n$ e $\boldsymbol{y}^i \in \mathbb{R}^n$, $i \in \{1,\ldots,m\}$.

Funciones cóncavas y convexas

Ejercicio 21. Dados $\boldsymbol{d} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$ y $\varepsilon > 0$, se llama cono de Bishop-Phelps al conjunto

$$K(\boldsymbol{d}, \varepsilon) = \{\boldsymbol{x} \in \mathbb{R}^3 : \varepsilon \|\boldsymbol{d}\| \|\boldsymbol{x}\| \leq \langle \boldsymbol{x}, \boldsymbol{d} \rangle\}$$

Demuestre que $K(\boldsymbol{d}, \varepsilon)$ es convexo para todo $\boldsymbol{d} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$ y $\varepsilon > 0$.

Ejercicio 22. Dados $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$ y $\gamma \in [0, 1]$ se llama pétalo de Penot al conjunto

$$P_\gamma(\boldsymbol{a}, \boldsymbol{b}) = \{\boldsymbol{x} \in \mathbb{R}^2 : \gamma \|\boldsymbol{a} - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{b}\| \leq \|\boldsymbol{b} - \boldsymbol{a}\|\}$$

Demuestre que $P_\gamma(\boldsymbol{d}, \varepsilon)$ es convexo para todo $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$ y $\gamma \in [0, 1]$.

Ejercicio 23. Sea $g : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Se define $f(\boldsymbol{x}) = \exp[g(\boldsymbol{x})]$. Demuestre que f es convexa.

Extremos restringidos

Ejercicio 24. De un cartón de $20m^2$ se va a construir una caja rectangular sin tapa. Determine el máximo volumen de la caja utilizando multiplicadores de Lagrange.

Ejercicio 25. Determinar el paralelepípedo de lados paralelos a los ejes de coordenadas, y de máximo volumen que cabe dentro del elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Ejercicio 26. Para p, q, r números racionales positivos, considere la función $f(x, y, z) = x^p y^q z^r$ y determine el mayor valor de $f(x, y, z)$ cuando $x + y + z = a$, $x > 0$, $y > 0$ y $z > 0$.

Ejercicio 27. Determine todos los valores extremos de $f(x, y) = x^2 + y^2 + z$ en la región $x^2 + y^2 + z^2 \leq 1$, $z \geq 0$.

Ejercicio 28. Encontrar los puntos de la esfera

$$3x^2 + 3y^2 + 3z^2 = 18$$

que están más lejos y más cerca del punto $(3, 1, -1)$ mediante multiplicadores de Lagrange.

Ejercicio 29. Encuentre el máximo valor de la función

$$x + 2y + 3z$$

en la curva de la intersección del plano $2x - y + z = 1$ y el cilindro $x^2 + y^2 = 1$ mediante multiplicadores de Lagrange.

Ejercicio 30. Sean $p, q > 0$. Encontrar el mínimo de la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = \frac{x^p}{p} + \frac{y^q}{q}$$

sujeto a $x, y > 0$ y $xy = 1$. Usar este resultado para deducir la siguiente desigualdad cuando $\frac{1}{p} + \frac{1}{q} = 1$ y $x, y > 0$

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

Ejercicio 31. Dada la función $f : \mathbb{R}^3 \to \mathbb{R}$ tal que $f(x, y, z) = xy + z$, determine si existe mínimos y máximos globales de f en la región $A = \left\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 - xy + z^2 \leq 1\right\}$. En caso de existir, calcúlelos.

Ejercicio 32. Calcular los extremos y los puntos silla para las funciones

1. $f(x, y, z) = x^2 + y^2 + z^2$ tal que $xy = 1$.
2. $f(x, y, z) = x^2 + 3y^2 + 2z^2 - 2xy + 2xz$ tal que $x + y + z = 1$.

Ejercicio 33. Resuelva el siguiente problema:

$$\begin{aligned} \min_{x,y} \quad & x^2 - y^2 \\ \text{s.a} \quad & x^2 + y^2 = 1 \end{aligned}$$

Ejercicio 34. Resuelva el siguiente problema:

$$\begin{aligned} \min_{x,y,z} \quad & x + y + z \\ \text{s.a} \quad & x^2 + y^2 = 2 \\ & x + z = 1 \end{aligned}$$

Ejercicio 35. Resuelva el siguiente problema:

$$\begin{aligned} \min_{\boldsymbol{d}} \quad & \|\boldsymbol{c} - \boldsymbol{d}\|^2 \\ \text{s.a} \quad & A\boldsymbol{d} = 0 \end{aligned}$$

donde $c, d \in \mathbb{R}^n$, $A \in \mathcal{M}_{m \times n}$ y $m < n$.

1. Encuentre una solución $\boldsymbol{d}_0$ para el problema, caracterícela.
2. Determine las condiciones para que la solución encontrada sea única.
3. Demuestre que $\boldsymbol{d}_0$ satisface $\boldsymbol{c} \cdot \boldsymbol{d}_0 \geq 0$, y que

$$\boldsymbol{c} \cdot \boldsymbol{d}_0 \geq 0 \iff \exists \lambda \in \mathbb{R}^m \text{ tal que } A^T \lambda = \boldsymbol{c}$$

Ejercicio 36. Sean $a, b, c \in \mathbb{R}_{++}$. Considere la función

$$f(x, y, z) = x^a y^b z^c$$

Determine el mayor valor de $f(x, y, z)$ mediante multiplicadores de Lagrange cuando

$$x + y + z = a \quad , \quad x, y, z > 0$$

Ejercicio 37. Encuentre el mínimo de la función

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

sujeto a la restricción

$$x^2 + y^2 = 2$$

Ejercicio 38. Doña Paipa es una vendedora que tiene un carrito de sopaipillas a la entrada de Beaucheff y desea maximizar sus utilidades (ingreso menos costos) de la venta diaria. Lo que se sabe es:

1. La producción de sopaipillas es representable por medio de la función $f : \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x, y, z) = \sqrt{xyz}$ y (x, y, z) son los insumos aceite (x), masa (y) y gas (z).
2. Lo que interesa como horizonte de tiempo es la producción de un día. Cada día se pueden producir a lo más 500 sopaipillas.
3. El costo de los insumos es 15, 20 y 35 pesos respectivamente. Cada sopaipilla se vende a 100 pesos.

En base a esta información:

1. Plantee el problema de optimización.
2. Resuelva utilizando multiplicadores de Lagrange para obtener la cantidad óptima que debe vender.
3. Determine el margen de ganancia que tiene Doña Paipa luego de un día de trabajo.

Ejercicio 39. Suponga que ahora Doña Paipa logra instalar un local de pizzas frente al edificio CEC. Nuevamente, lo que le interesa es maximizar las utilidades de la venta diaria. La producción de pizzas es representable por medio de la función $f : \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x, y, z) = (x^2 + y^2 + z^2 + w^2)^{1/2}$ y (x, y, z, w) son los insumos masa (x), queso (y), otro ingrediente (z) y electricidad (w).
Ahora los costos son distintos y todos los insumos tienen un costo unitario de 100 pesos. Cada pizza se vende a 500 pesos y se pueden producir a lo más 200 pizzas por día.

En base a esta información plantee el problema de optimización y resuelva.

Ejercicio 40. La viña Don Pachá produce tres tipos de vinos (Carménère, Cavernet Sauvignon y Merlot). Por razones enológicas que no son parte del apunte la producción de cada vino viene dada por combinaciones de capital (K) y mano de obra (L) representables por medio de las siguientes funciones

1. Carmenere: $f(K, L) = 0{,}8K^{0{,}3}L^{0{,}7}$
2. Cavernet Sauvignon: $f(K, L) = KL$
3. Merlot: $f(K, L) = (K^{1/2} + L^{1/2})^2$

Si el costo de una unidad de capital es m y de una unidad mano de obra es n. Obtenga la demanda por factores proveniente de la minimización de costos para cada caso, planteando el problema de minimización general y resuelva utilizando multiplicadores de Lagrange.

Ejercicio 41. Considere la forma cuadrática

$$(x, y) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

minimice su valor sujeto a la restricción $\|(x, y)\| = 1$.

Criterio de 2^do^ orden para extremos restringidos

Ejercicio 42. Sea $f : \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x, y, z) = \ln(xyz^3)$

1. Encuentre el valor máximo de f sobre

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 5r^2,\ x > 0,\ y > 0,\ z > 0\}$$

   INDICACIÓN. Verifique que el hessiano del lagrangeano es definido positivo sobre dicho punto.

2. Usando lo anterior, muestre que para números reales positivos se cumple

$$abc^3 \leq 27\left(\frac{a+b+c}{5}\right)^5$$

Ejercicio 43. Sea $Q \in \mathcal{M}_{n\times n}(\mathbb{R})$, $\boldsymbol{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$ y sea $f(\boldsymbol{x}) = \boldsymbol{x}^T Q\boldsymbol{x} + \boldsymbol{b}\cdot\boldsymbol{x} + c$. Encuentre las condiciones sobre Q para la existencia de máximos y mínimos para f.

Además, encuentre las mismas condiciones de la parte anterior pero bajo la restricción

$$\sum_{i=1}^{n} \frac{x_i^2}{n} = 1$$

CAPÍTULO 3

Integración

Nos interesa extender la noción de "área bajo una curva", formalizada por la integral de Riemann en una variable, a la de "área bajo una superficie" en $\mathbb{R}^n$. Luego estudiaremos las propiedades fundamentales de la integral de Riemann en varias variables. Finalmente veremos algunas aplicaciones a problemas físicos.

3.1. Integral de Riemann en $\mathbb{R}^2$

3.1.1. Definiciones

Definición 3.1. $R \subseteq \mathbb{R}^2$ es un rectángulo si y sólo si $R = [a_1, b_1] \times [a_2, b_2]$ con $a_1, b_1, a_2, b_2 \in \mathbb{R}$.

El área de un rectángulo $R = [a_1, b_1] \times [a_2, b_2]$ se define como

$$V(R) = (b_1 - a_1)(b_2 - a_2)$$

Definición 3.2. Para $m \in \mathbb{N}$, definimos la m-equipartición del intervalo $[a, b]$, con $a, b \in \mathbb{R}$, como

$$\{[c_0, c_1], [c_1, c_2], \ldots, [c_{m-1}, c_m]\}$$

donde los puntos c_i vienen dados por: $c_i = a + i\frac{b-a}{m}$, para $i = 0, \ldots, m$.

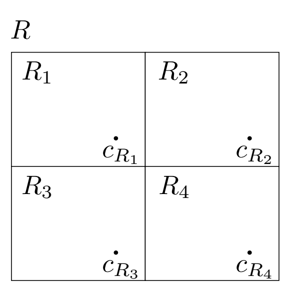


Figura 3.1: 2-equipartición y selección

Definición 3.3. Para $m \in \mathbb{N}$, definimos la m-equipartición del rectángulo $R = [a_1, b_1] \times [a_2, b_2]$ como $P = \{I_1 \times I_2 : I_i \in P_i, i = 1, 2\}$ donde P_i es la m-equipartición del intervalo $[a_i, b_i]$ para $i = 1, 2$. Denotaremos la m-equipartición de R como $P_m(R)$.

Definición 3.4. Dado un rectángulo $R \subseteq \mathbb{R}^2$ y $m \in \mathbb{N}$, decimos que $(\boldsymbol{c}_P)_{P \in P_m(R)}$ es una selección (para $P_m(R)$) si $(\forall P \in P_m(R))$ se tiene $\boldsymbol{c}_P \in P$.

Notar que cada elemento de una m-equipartición es un rectángulo y que una m-equipartición es finita, luego la siguiente definición tiene sentido:

Definición 3.5. Sea $m \in \mathbb{N}$, $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$ y una selección $(\boldsymbol{c}_P)_{P \in P_m(R)}$. Se define la suma de Riemann asociada a f y $(\boldsymbol{c}_P)_{P \in P_m(R)}$ como:

$$S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) = \sum_{P \in P_m(R)} f(\boldsymbol{c}_P) V(P)$$

Definición 3.6. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$. Decimos que f es Riemann integrable en R si y sólo si $(\exists S \in \mathbb{R})(\forall \varepsilon > 0)(\exists m_0 \in \mathbb{N})(\forall m \geq m_0)$

$$\left| S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) - S \right| < \varepsilon$$

para cualquier selección $(\boldsymbol{c}_P)_{P \in P_m(R)}$.

S se llama integral (de Riemann) de f sobre R y se denota:

$$\int_R f$$

La integral de f también se anota:

$$\int_R f(\boldsymbol{x})\, d\boldsymbol{x}$$

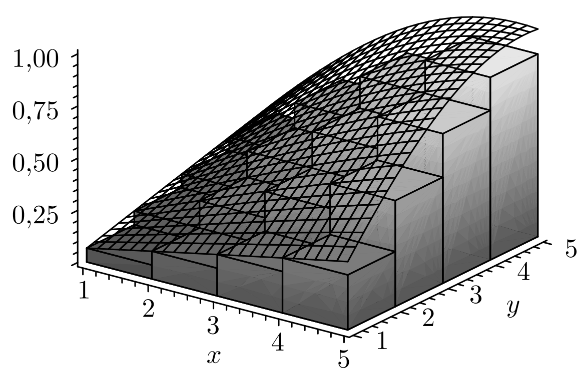


Figura 3.2: Suma de Riemann para la función $f(x, y) = \operatorname{sen}\left(\frac{xy}{15}\right)$ para $P_4([1, 5]^2)$

Ejemplo 3.1. Para $R = [0, 1]^2$, no es integrable en R la función $f : R \to \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} 1 & \text{si } (x, y) \in \mathbb{Q} \times \mathbb{Q} \\ 0 & \text{si } (x, y) \notin \mathbb{Q} \times \mathbb{Q} \end{cases}$$

3.1.2. Propiedades básicas

Proposición 3.1. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$. Si f es integrable en R entonces f es acotada en R.

Demostración. En efecto, sea $\varepsilon = 1$, $m = m_0$ en la definición de integrabilidad. Luego

$$\left| S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) - S\right| < 1$$
$$\left| \sum_{P \in P_m(R)} f(\boldsymbol{c}_P) V(P) - S\right| < 1$$
$$\left| \sum_{P \in P_m(R)} f(\boldsymbol{c}_P) V(P)\right| < 1 + |S|$$

y para un cierto $P_0 \in P_m(R)$

$$\left|f(c_{P_0})\right| V(P_0) < 1 + |S| + \left| \sum_{P \in P_m(R) \setminus \{P_0\}} f(\boldsymbol{c}_P) V(P)\right|$$
$$\left|f(c_{P_0})\right| < \frac{1}{V(P_0)} \left(1 + |S| + \left| \sum_{P \in P_m(R) \setminus \{P_0\}} f(\boldsymbol{c}_P) V(P)\right|\right)$$

Fijando los $\boldsymbol{c}_P$, para $P \in P_m(R) \setminus \{P_0\}$ y notando que c_{P_0} es arbitrario en P_0 se concluye que f es acotada en P_0. Como además P_0 es arbitrario en $P_m(R)$ se concluye que f es acotada en cada $P \in P_m(R)$, luego f es acotada en R. ■

La siguiente propiedad es análoga a la condición de Cauchy para una sucesión, luego es útil por ejemplo para estudiar la integrabilidad de una función cuando no se conoce el valor de su integral.

Proposición 3.2. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$. Son equivalentes:

1. $\forall \varepsilon > 0, \exists m_0 \in \mathbb{N}, \forall k, m \geq m_0, \forall (\boldsymbol{c}'_P)_{P \in P_k(R)}$ selección, $\forall (\boldsymbol{c}_P)_{P \in P_m(R)}$ selección
$$\left| S\left(f, (\boldsymbol{c}'_P)_{P \in P_k(R)}\right) - S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right)\right| < \varepsilon$$
2. f es integrable en R.

Demostración.
($\Rightarrow$): Fijemos ciertas selecciones $(\boldsymbol{c}^m_P)_{P \in P_m(R)}$, para cada $m \in \mathbb{N}$. Luego, por la hipótesis, la sucesión $\left(S\left(f, (\boldsymbol{c}^m_P)_{P \in P_m(R)}\right)\right)_{m \in \mathbb{N}}$ resulta ser de Cauchy y converge a un cierto $S \in \mathbb{R}$. Sea $\varepsilon > 0$. Sea $m_1 \geq m_0$ tal que $(\forall m \geq m_0)$

$$\left| S\left(f, (\boldsymbol{c}^m_P)_{P \in P_m(R)}\right) - S\right| < \varepsilon$$

Sea $m \geq m_0$ y una selección $(\boldsymbol{c}_P)_{P \in P_m(R)}$ arbitraria. Así, nuevamente con la hipótesis, resulta que

$$\begin{aligned}\left| S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) - S\right| &\leq \left| S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) - S\left(f, (\boldsymbol{c}^{m_1}_P)_{P \in P_{m_1}}\right)\right| \\ &\quad + \left| S\left(f, (\boldsymbol{c}^{m_1}_P)_{P \in P_{m_1}}\right) - S\right| \\ &< 2\varepsilon\end{aligned}$$

($\Leftarrow$): Sea $\varepsilon > 0$, sea $m_0 \in \mathbb{N}$ tal que $(\forall m \geq m_0)$

$$\left| S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - S \right| < \varepsilon$$

para toda elección de los $(\boldsymbol{c}_P)_{P\in P_m(R)}$. Sean $k, m \geq m_0$ arbitrarios, sean $(\boldsymbol{c}_P)_{P\in P_m(R)}$, $(\boldsymbol{c}'_P)_{P\in P_k(R)}$ selecciones arbitrarias. Luego:

$$\begin{aligned}\left| S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - S\left(f, (\boldsymbol{c}'_P)_{P\in P_k(R)}\right) \right| &\leq \left| S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - S \right| \\ &\quad + \left| S - S\left(f, (\boldsymbol{c}'_P)_{P\in P_k(R)}\right) \right| \\ &\leq 2\varepsilon\end{aligned}$$

■

El siguiente lema es una consecuencia de la proposición 3.2 y nos da una condición necesaria y suficiente de integrabilidad más fácil de verificar en varias de las propiedades que siguen. La principal diferencia con la proposición 3.2 es que en vez de comparar todos los pares de particiones, compara pares de particiones que son una un refinamiento de la otra.

Lema 3.1. *Sea $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$. Son equivalentes*

1. $(\forall\varepsilon > 0)(\exists m_0 \in \mathbb{N})(\forall k \in \mathbb{N})(\forall m \geq m_0)(\forall(\boldsymbol{c}_P)_{P\in P_m(R)}, (\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ *selección*)

$$\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P) \right| V(Q) < \varepsilon$$

2. *f es integrable en R.*

Demostración. Usaremos la proposición 3.2 para sustituir la condición f integrable en R.

($\Rightarrow$): Sea $\varepsilon > 0$. Por hipótesis

$$(\exists m_0 \in \mathbb{N})(\forall k \in \mathbb{N})(\forall m \geq m_0)(\forall(\boldsymbol{c}_P)_{P\in P_m(R)} \text{ selección})(\forall(\boldsymbol{c}'_P)_{P\in P_{km}(R)} \text{ selección})$$

$$\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P) \right| V(Q) < \varepsilon$$

Sean $k, m \geq m_0$, sean $(\boldsymbol{c}_P)_{P\in P_m(R)}$, $(\boldsymbol{c}'_P)_{P\in P_k(R)}$ selecciones. Sea $(\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ selección. Luego:

$$\begin{aligned}&\left| S\left(f, (\boldsymbol{c}'_Q)_{Q\in P_{km}(R)}\right) - S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) \right| \\ &= \left| \sum_{Q\in P_{km}(R)} f(\boldsymbol{c}'_Q)V(Q) - \sum_{P\in P_m(R)} f(\boldsymbol{c}_P)V(P) \right| \\ &= \left| \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} f(\boldsymbol{c}'_Q)V(Q) - \sum_{P\in P_m(R)} f(\boldsymbol{c}_P) \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} V(Q) \right| \\ &\leq \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P) \right| V(Q) \\ &< \varepsilon\end{aligned}$$

y análogamente

$$\left|S\left(f,(\boldsymbol{c}'_Q)_{Q\in P_{km}(R)}\right)-S\left(f,(\boldsymbol{c}'_P)_{P\in P_k(R)}\right)\right|<\varepsilon$$

Así

$$\begin{aligned}&\left|S\left(f,(\boldsymbol{c}_P)_{P\in P_m(R)}\right)-S\left(f,(\boldsymbol{c}'_P)_{P\in P_k(R)}\right)\right|\\ &\leq\left|S\left(f,(\boldsymbol{c}_P)_{P\in P_m(R)}\right)-S\left(f,(\boldsymbol{c}'_Q)_{Q\in P_{km}(R)}\right)\right|\\ &\quad+\left|S\left(f,(\boldsymbol{c}'_Q)_{Q\in P_{km}(R)}\right)-S\left(f,(\boldsymbol{c}'_P)_{P\in P_k(R)}\right)\right|\\ &<2\varepsilon\end{aligned}$$

($\Leftarrow$): Sea $\varepsilon>0$. Por hipótesis, $(\exists m_0\in\mathbb{N})(\forall k,m\geq m_0)(\forall(\boldsymbol{c}'_P)_{P\in P_k(R)}$ selección$)(\forall(\boldsymbol{c}_P)_{P\in P_m(R)}$ selección$)$

$$\left|S\left(f,(\boldsymbol{c}'_P)_{P\in P_k(R)}\right)-S\left(f,(\boldsymbol{c}_P)_{P\in P_m(R)}\right)\right|<\varepsilon$$

Sean $k\in\mathbb{N}$, $m\geq m_0$, sean $(\boldsymbol{c}_P)_{P\in P_m(R)},(\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ selecciones. Para $P\in P_m(R)$ definamos $\boldsymbol{c}_P^+=\text{argmax}\{f(\boldsymbol{c}'_Q):Q\in P_{km}(R),Q\subseteq P\}\cup\{f(\boldsymbol{c}_P)\}$, $\boldsymbol{c}_P^-=\text{argmin}\{f(\boldsymbol{c}'_Q):Q\in P_{km}(R),Q\subseteq P\}\cup\{f(\boldsymbol{c}_P)\}$. Luego

$$\begin{aligned}&\sum_{P\in P_m(R)}\sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}}\left|f(\boldsymbol{c}'_Q)-f(\boldsymbol{c}_P)\right|V(Q)\\ &\leq\sum_{P\in P_m(R)}\sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}}\left(f(\boldsymbol{c}_P^+)-f(\boldsymbol{c}_P^-)\right)V(Q)\\ &=\sum_{P\in P_m(R)}\left(f(\boldsymbol{c}_P^+)-f(\boldsymbol{c}_P^-)\right)V(P)\\ &=S\left(f,(\boldsymbol{c}_Q^+)_{Q\in P_m(R)}\right)-S\left(f,(\boldsymbol{c}_P^-)_{P\in P_m(R)}\right)\\ &<\varepsilon\end{aligned}$$

■

Posteriormente (teorema 4.12) probaremos que toda función continua sobre un rectángulo es uniformemente continua en él. En realidad lo probaremos para conjuntos mucho más generales que rectángulos, llamados compactos. Esta propiedad se utiliza en la demostración de la siguiente proposición.

Proposición 3.3. Sea $R\subseteq\mathbb{R}^2$ rectángulo, $f:R\to\mathbb{R}$. Si f es continua en R entonces f es integrable en R.

Demostración. Como R es compacto y f es continua en R se tiene que f es uniformemente continua en R. Luego, para $\varepsilon>0$ arbitrario, $(\exists\delta>0)(\forall\boldsymbol{x},\boldsymbol{y}\in R)$

$$\|\boldsymbol{x}-\boldsymbol{y}\|<\delta\Rightarrow\left|f(\boldsymbol{x})-f(\boldsymbol{y})\right|<\varepsilon$$

Recurriremos al lema 3.1. Sea $m_0\in\mathbb{N}$ tal que $(\forall P\in P_{m_0})(\forall\boldsymbol{x},\boldsymbol{y}\in P)$ $\|\boldsymbol{x}-\boldsymbol{y}\|<\delta$. Sea $m\geq m_0$, $k\in\mathbb{N}$. Sean $(\boldsymbol{c}_P)_{P\in P_m(R)}$, $(\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ selecciones.

Por la uniforme continuidad se tiene que

$$\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left|f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q) \leq \varepsilon V(R)$$

■

Proposición 3.4. (Propiedades de la clase de funciones integrables)
Sean $R \subseteq \mathbb{R}^2$ rectángulo, $f, g : R \to \mathbb{R}$ funciones integrables y $c \in \mathbb{R}$. Se tienen las siguientes propiedades:

1. Linealidad: $f + cg$ es integrable en R, entonces

$$\int_R f + cg = \int_R f + c\int_R g$$

2. Monotonía: si $(\forall x \in R) f(x) \leq g(x)$, entonces

$$\int_R f \leq \int_R g$$

3. $|f|$ es integrable en R, entonces

$$\left|\int_R f\right| \leq \int_R |f|$$

Demostración.

1. Sea $\varepsilon > 0$ arbitrario. Sea $m_0 \in \mathbb{N}$ tal que $(\forall m \geq m_0)(\forall (\boldsymbol{c}_P)_{P\in P_m(R)}$ selección$)$

$$\left|S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - \int_R f\right| < \varepsilon$$

y

$$\left|S\left(g, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - \int_R f\right| < \varepsilon$$

Por otra parte, se tiene:

$$S\left(f + cg, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) = S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) + cS\left(g, (\boldsymbol{c}_P)_{P\in P_m(R)}\right)$$

Así,

$$\begin{aligned}
&\left|S\left(f + cg, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - (\int_R f + c\int_R g)\right| \\
&\leq \left|S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - \int_R f\right| + |c| \left|S\left(g, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - \int_R g\right| \\
&\leq \left(1 + |c|\right)\varepsilon
\end{aligned}$$

Luego, por definición de integral de Riemann, se concluye.

2. Es directo de considerar que en este caso se cumple que $\forall m \in \mathbb{N}$, $\forall (\boldsymbol{c}_P)_{P\in P_m(R)}$ selección

$$S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) \leq S\left(g, (\boldsymbol{c}_P)_{P\in P_m(R)}\right)$$

y aplicar la definición de integral de Riemann.

3. Veamos que $|f|$ es integrable en R por medio del lema 3.1.

En efecto, sea $\varepsilon > 0$. Como f es integrable, por el mismo lema $(\exists m_0 \in \mathbb{N})(\forall k \in \mathbb{N})(\forall m \geq m_0)(\forall (\boldsymbol{c}_P)_{P\in P_m(R)}, (\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ selección)

$$\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q) < \varepsilon$$

Luego, sean $k \in \mathbb{N}$, $m \geq m_0$. Se sabe que $(\forall x, y \in \mathbb{R})$

$$\big|\, |x| - |y| \,\big| \leq |x - y|$$

Así

$$\begin{aligned}
\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} &\big| \,|f|\,(\boldsymbol{c}'_Q) - |f|\,(\boldsymbol{c}_P)\big| V(Q) \\
&\leq \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q) \\
&\leq \varepsilon
\end{aligned}$$

En conclusión, $|f|$ es integrable en R. Finalmente, se tiene

$$-\,|f| \leq f \leq |f|$$

que, con la parte anterior, implica que

$$-\int_R |f| \leq \int_R f \leq \int_R |f|$$

es decir

$$\left| \int_R f \right| \leq \int_R |f|$$

■

Proposición 3.5. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$. Si f es integrable en R entonces

$$V(R) \inf_{\boldsymbol{x}\in R} f(\boldsymbol{x}) \leq \int_R f \leq V(R) \sup_{\boldsymbol{x}\in R} f(\boldsymbol{x})$$

Demostración. Basta notar que $\forall m \in \mathbb{N}$, $\forall (\boldsymbol{c}_P)_{P\in P_m(R)}$ selección

$$V(R) \inf_{\boldsymbol{x}\in R} f(\boldsymbol{x}) \leq S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) \leq V(R) \sup_{\boldsymbol{x}\in R} f(\boldsymbol{x})$$

■

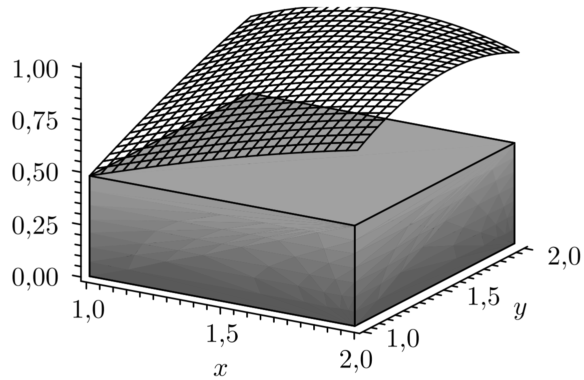


Figura 3.3: Ejemplo de que $V(R)\inf_{\boldsymbol{x}\in R} f(\boldsymbol{x}) \leq \int_R f$

3.1.3. Integración de sucesiones de funciones

Proposición 3.6. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $(f_k)_{k\in\mathbb{N}}$ sucesión de funciones, $f_k : R \to \mathbb{R}$, que convergen uniformemente en R a $f : R \to \mathbb{R}$. Entonces f es integrable en R y

$$\int_R f = \lim_{k\to\infty} \int_R f_k$$

Demostración. Veamos que f es integrable en R por medio del lema 3.1. Sea $\varepsilon > 0$. Sea $k_0 \in \mathbb{N}$ tal que

$$\sup_{x\in R} \left| f(\boldsymbol{x}) - f_{k_0}(\boldsymbol{x}) \right| < \varepsilon \tag{3.1}$$

De acuerdo al lema 3.1, sea $m_0 \in \mathbb{N}$ tal que
$(\forall k \in \mathbb{N})(\forall m \geq m_0)(\forall (\boldsymbol{c}_P)_{P\in P_m(R)}, (\boldsymbol{c}'_P)_{P\in P_{km}(R)}$ selección)

$$\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f_{k_0}(\boldsymbol{c}'_Q) - f_{k_0}(\boldsymbol{c}_P) \right| V(Q) < \varepsilon \tag{3.2}$$

Luego, sean $k \in \mathbb{N}$, $m \geq m_0$. Por la desigualdad triangular se tiene que

$$\begin{aligned}
&\sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P) \right| V(Q) \\
&\qquad \leq \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f(\boldsymbol{c}'_Q) - f_{k_0}(c'_Q) \right| V(Q) \\
&\qquad\quad + \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f_{k_0}(c'_Q) - f_{k_0}(\boldsymbol{c}_P) \right| V(Q) \\
&\qquad\quad + \sum_{P\in P_m(R)} \sum_{\substack{Q\in P_{km}(R)\\ Q\subseteq P}} \left| f_{k_0}(\boldsymbol{c}_P) - f(\boldsymbol{c}_P) \right| V(Q)
\end{aligned}$$

y aplicando 3.1 y 3.2 se obtiene

$$\leq \varepsilon\big(2V(R)+1\big)$$

En conclusión, f es integrable en R. Por otra parte, por la proposición 3.5

$$\begin{aligned}\left|\int_R f_k - \int_R f\right| &= \left|\int_R (f_k - f)\right| \\ &\leq \int_R |f_k - f| \\ &\leq V(R)\sup_{x\in R}|f_k(x) - f(x)|\end{aligned}$$

lo cual implica que

$$\int_R f = \lim_{k\to\infty}\int_R f_k$$

■

Ejemplo 3.2. Una sucesión de funciones integrables que converge puntualmente a un límite que no lo es. Sea la numeración $\{q_0, q_1, \dots\} = [0,1]\cap\mathbb{Q}$, sea $I = [0,1]$ y las funciones $f, f_n : I \to \mathbb{R}$ dadas por

$$f_n(\boldsymbol{x}) = \begin{cases} 1 & \text{si } (x,y)\in\{q_1,\dots,q_n\} \\ 0 & \text{si } (x,y)\notin\{q_1,\dots,q_n\}\end{cases}$$

$$f(\boldsymbol{x}) = \begin{cases} 1 & (x,y)\in[0,1]\cap\mathbb{Q} \\ 0 & (x,y)\notin[0,1]\cap\mathbb{Q}\end{cases}$$

En I, se tiene que f_n converge a f puntualmente (pero no uniformemente) y cada f_n es integrable, pero f no es integrable.

3.1.4. Extensión de la clase de funciones integrables

Aún cuando la clase de funciones continuas es muy amplia, todavía no es suficiente para muchas aplicaciones. Así, a continuación veremos una condición suficiente de integrabilidad un poco más débil que la continuidad, esto es, que una función sea continua salvo sobre la unión de grafos de funciones continuas.

Recordar que el grafo de una función $g : [a,b] \to \mathbb{R}$ es

$$\text{grafo}(g) = \big\{(x, g(x)) \in \mathbb{R}^2 : x \in [a,b]\big\}$$

Intuitivamente, el lema siguiente nos dice que el grafo de una función continua tiene "volumen cero".

Lema 3.2. *Sea $f : [a,b] \to [c,d]$ continua. Denotemos $R = [a,b]\times[c,d]$. Entonces*

$$\lim_{m\to\infty}\sum_{\substack{P\in P_m(R)\\ P\cap\text{grafo}(g)\neq\varnothing}} V(P) = 0$$

Demostración. Sea $\varepsilon > 0$. Como f continua en $[a,b]$ compacto, se tiene que f es uniformemente continua en $[a,b]$ (teorema 4.12). Luego $(\exists\delta > 0)(\forall x, y \in [a,b])$

$$|x - y| < \delta \Rightarrow \big|f(x) - f(y)\big| < \varepsilon$$

Así, escojamos $m_0 \in \mathbb{N}$ tal que $(b-a)(d-c)/m_0 \leq \varepsilon$ y

$$(b-a)/m_0 < \delta \tag{3.3}$$

Entonces, se cumple que $(\forall m \geq m_0)$

$$\sum_{\substack{P\in P_m(R)\\ P\cap \text{grafo}(g)\neq\varnothing}} V(P) = \frac{(b-a)(d-c)}{m^2}\big|\{P \in P_m(R) : P \cap \text{grafo}(g) \neq \varnothing\}\big|$$

Gracias a la uniforme continuidad y (3.3)

$$\big|\{P \in P_m(R) : P \cap \text{grafo}(g) \neq \varnothing\}\big| \leq m(\frac{\varepsilon}{(d-c)/m} + 1)$$

Así, $\forall m \geq m_0$

$$\begin{aligned}\sum_{\substack{P\in P_m(R)\\ P\cap \text{grafo}(g)\neq\varnothing}} V(P) &\leq \frac{(b-a)(d-c)}{m^2} m(\frac{\varepsilon}{(d-c)/m} + 1)\\ &= \varepsilon(b-a) + \frac{(b-a)(d-c)}{m}\\ &\leq \varepsilon(b-a+1)\end{aligned}$$

■

Para $A \subseteq \mathbb{R}^n$ acotado, se define el diámetro de A como

$$\text{diam}(A) = \sup_{x,y\in A} \|\boldsymbol{x} - \boldsymbol{y}\|$$

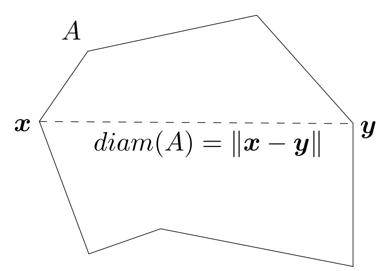


Figura 3.4: Diámetro de un conjunto $A \subseteq \mathbb{R}^2$

Proposición 3.7. Sea $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$ acotada en R y continua en $R \to \text{grafo}(g)$, donde $g : [a,b] \to [c,d]$ es una función continua. Entonces f es integrable en R.

Demostración. Sea $K \in \mathbb{R}$ tal que $(\forall x \in R)$

$$\left|f(\boldsymbol{x})\right| \leq K \tag{3.4}$$

Recurriremos al lema 3.1. Sea $\varepsilon > 0$ arbitrario. De acuerdo al lema 3.2, sea $m_0 \in \mathbb{N}$ tal que $(\forall m \geq m_0)$

$$\sum_{\substack{P \in P_m(R) \\ P \cap \text{grafo}(g) \neq \varnothing}} V(P) \leq \varepsilon \tag{3.5}$$

Denotemos $P_m^* = \{P \in P_m(R) : P \cap \text{grafo}(g) = \varnothing\}$, $R_m^* = \bigcup_{P \in P_m^*} P$. Como $R_{m_0}^*$ es compacto y f es continua en $R_{m_0}^*$ se tiene que f es uniformemente continua en $R_{m_0}^*$. Luego $(\exists \delta > 0)(\forall \boldsymbol{x}, \boldsymbol{y} \in R_{m_0}^*)$

$$\|\boldsymbol{x} - \boldsymbol{y}\| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon \tag{3.6}$$

Sea $m_1 \geq m_0$ tal que $(\forall P \in P_{m_1}(R))\, \text{diam}(P) < \delta$. Sea $m \geq m_1$, $k \in \mathbb{N}$. Sean $(\boldsymbol{c}_P)_{P \in P_m(R)}$, $(\boldsymbol{c}'_P)_{P \in P_{km}(R)}$ selecciones. Así

$$\begin{aligned}
&\sum_{P \in P_m(R)} \sum_{\substack{Q \in P_{km}(R) \\ Q \subseteq P}} \left|f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q) \\
&\quad = \sum_{\substack{P \in P_m(R) \\ P \subseteq R_{m_0}^*}} \sum_{\substack{Q \in P_{km}(R) \\ Q \subseteq P}} \left|f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q) + \sum_{\substack{P \in P_m(R) \\ P \nsubseteq R_{m_0}^*}} \sum_{\substack{Q \in P_{km}(R) \\ Q \subseteq P}} \left|f(\boldsymbol{c}'_Q) - f(\boldsymbol{c}_P)\right| V(Q)
\end{aligned}$$

que, con (3.4) y (3.6), implica que

$$\leq \varepsilon V(R) + 2K \sum_{\substack{P \in P_m(R) \\ P \nsubseteq R_{m_0}^*}} V(P)$$

y con (3.5) resulta

$$\leq \varepsilon V(R) + \varepsilon 18K$$

■

Corolario 3.1. *Sea $R = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ rectángulo, $f : R \to \mathbb{R}$ acotada en R y continua en $R \to \bigcup_{i=1}^n \text{grafo}(g_i)$, donde $g_i : [a, b] \to [c, d]$ es continua. Entonces f es integrable en R.*

3.1.5. Teorema de Fubini

El siguiente teorema expresa la integral de una función en 2 variables como la aplicación iterada de 2 integrales en una variable bajo hipótesis mínimas y permite escoger arbitrariamente el orden de estas integrales en una variable cuando la función a integrar es continua.

La demostración la haremos en la sección 3.2 en $\mathbb{R}^n$.

Teorema 3.1. (Teorema de Fubini en $\mathbb{R}^2$)

Sea $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$ *rectángulo,* $f : R \to \mathbb{R}$.

1. *Si* f *integrable en* R *y* $(\forall x \in [a,b])$ $f(x,\cdot)$ *integrable en* $[c,d]$*, entonces*

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

2. *Si* f *continua en* R*, entonces*

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

Ejemplo 3.3. Para $R = [0,1] \times [0,1]$, calcular

$$\int_R x^2 + y$$

Solución. Por el teorema de Fubini, caso continuo, se tiene que

$$\begin{aligned}
\int_R x^2 + y &= \int_0^1 \left(\int_0^1 x^2 + y\, dy \right) dx \\
&= \int_0^1 \left(x^2 y + \frac{y^2}{2}\bigg|_{y=0}^1 \right) dx \\
&= \int_0^1 x^2 + \frac{1}{2}\, dx \\
&= \frac{x^3}{3} + \frac{x}{2}\bigg|_0^1 = \frac{5}{6}
\end{aligned}$$

$\square$

Ejemplo 3.4. Veamos un caso en el que las integrales iteradas existen y son iguales, pero la función no es integrable.

Consideremos el cuadrado unitario $R = [0,1]^2 \subseteq \mathbb{R}^2$ y la sucesión de cuadrados $(R_k)_{k\in\mathbb{N}}$ contenidos en él dada por la figura 3.5. Dividamos cada R_k en 4 cuadrados iguales $R_k^{(1)}, R_k^{(2)}, R_k^{(3)}, R_k^{(4)}$. Definamos la función $f : R \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \frac{1}{V(R_k)} & \text{si } (x,y) \in \operatorname{int} R_k^{(1)} \cup \operatorname{int} R_k^{(3)} \\ -\frac{1}{V(R_k)} & \text{si } (x,y) \in \operatorname{int} R_k^{(2)} \cup \operatorname{int} R_k^{(4)} \\ 0 & \text{en otro caso} \end{cases}$$

Para cada $y \in [0,1]$ es claro que

$$\int_0^1 f(x,y)\, dx = 0$$

y análogamente para cada $x \in [0,1]$

$$\int_0^1 f(x,y)\, dy = 0$$

Luego, las integrales iteradas son ambas nulas, esto es:

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = \int_0^1 \int_0^1 f(x,y)\,dx\,dy = 0.$$

Para convencernos de que f no es integrable, es suficiente ver que $|f|$ no es integrable. En efecto, $|f|$ está dada por

$$|f|(x,y) = \begin{cases} \frac{1}{V(R_k)} & \text{si } (x,y) \in R_k \\ 0 & \text{en otro caso} \end{cases}$$

Si $|f|$ fuera integrable, como los R_k son disjuntos, se tendría que

$$\int_R |f| = \sum_{k \in \mathbb{N}} \int_{R_k} |f|$$

pero

$$\int_{R_k} |f| = 1$$

con lo que $|f|$ no puede ser integrable en R.

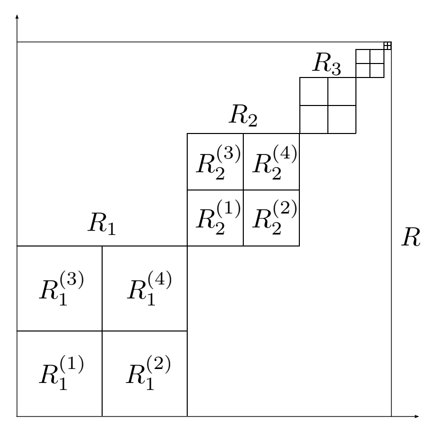


Figura 3.5: El dominio del ejemplo 3.4.

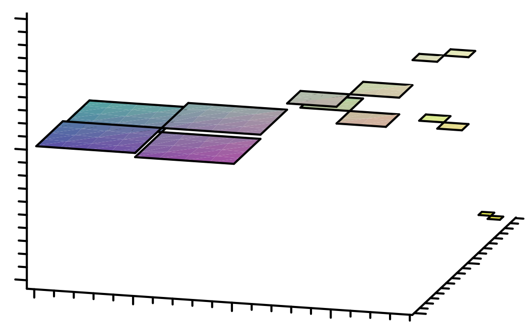


Figura 3.6: La función del ejemplo 3.4.

3.1.6. Integral en $\mathbb{R}^2$ sobre dominios generales

A continuación extenderemos la definición de integral para considerar la integración de funciones sobre algunos dominios un tanto más generales que los rectángulos.

Definición 3.7. Sea $R \subseteq \mathbb{R}^2$ rectángulo, $A \subseteq R$, $f : A \to \mathbb{R}$. Denotemos $f_0 : R \to \mathbb{R}$ a la función dada por, para $x \in R$:

$$f_0(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) & \text{si } \boldsymbol{x} \in A \\ 0 & \text{si } \boldsymbol{x} \notin A \end{cases}$$

Si f_0 es integrable sobre R entonces se dice que f es función integrable sobre A y:

$$\int_A f = \int_R f_0$$

Definición 3.8. La función indicatriz de un conjunto $A \subseteq \mathbb{R}^2$ se define:

$$\mathbf{1}_A(\boldsymbol{x}) = \begin{cases} 1 & \text{si } \boldsymbol{x} \in A \\ 0 & \text{si } \boldsymbol{x} \notin A \end{cases}$$

Definición 3.9. Decimos que $D \subseteq \mathbb{R}^2$ es

1. Dominio de tipo 1 si y sólo si $\exists a, b \in \mathbb{R}$, $\exists \phi_1, \phi_2 : [a, b] \to \mathbb{R}$ funciones continuas tales que $(\forall x \in [a, b])\phi_1(x) \leq \phi_2(x)$ y
$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$
2. Dominio de tipo 2 si y sólo si $\exists c, d \in \mathbb{R}$, $\exists \psi_1, \psi_2 : [c, d] \to \mathbb{R}$ funciones continuas tales que $(\forall x \in [c, d])\psi_1(x) \leq \psi_2(x)$ y
$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$
3. Dominio de tipo 3 si y sólo si D es de tipo 1 y 2,
4. Dominio elemental si y sólo si D es de tipo 1 ó 2.

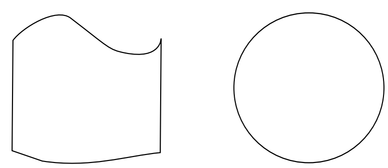


Figura 3.7: Dominio del tipo 1 que no es del tipo 2 y dominio del tipo 3.

Si $R \subseteq \mathbb{R}^2$ es un rectángulo y $D \subseteq R$ es una región elemental entonces f_0 (dada por la definición 3.7) es continua en R salvo sobre la unión finita de grafos de funciones. Luego $\int_D f$ existe. En general, supongamos que tenemos una función $f : A \subseteq R \to \mathbb{R}$ con R rectángulo y A dominio del tipo 1, digamos:
$$A = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}$$
con $a, b \in \mathbb{R}$, $\phi_1, \phi_2 : [a, b] \to \mathbb{R}$ funciones continuas. Entonces, gracias al teorema de Fubini (teorema 3.1), se tiene que:

$$\int_A f = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\, dy\, dx$$

Ejemplo 3.5. Calcular
$$\int_T (x^3 y + \cos(x))\, dy\, dx$$
donde
$$T = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq \pi/2, 0 \leq y \leq x\}$$
Solución. Notar que T es una región del tipo 1. Por definición se tiene que

$$\int_T (x^3 y + \cos(x))\, dy\, dx = \int_{[0,\pi/2]^2} \mathbf{1}_T (x^3 y + \cos(x))\, dy\, dx$$

y, en virtud del teorema de Fubini:

$$= \int_0^{\pi/2} \int_0^{\pi/2} \mathbf{1}_T \left(x^3 y + \cos(x)\right) \, dy \, dx$$

Más aún, con la definición de T:

$$\begin{aligned}
&= \int_0^{\pi/2} \int_0^{x} \left(x^3 y + \cos(x)\right) \, dy \, dx \\
&= \int_0^{\pi/2} \left(\frac{x^3 y^2}{2} + y\cos(x) \right) \Bigg|_{y=0}^{x} dx \\
&= \int_0^{\pi/2} \frac{x^5}{2} + x\cos(x) \, dx \\
&= \left(\frac{x^6}{12} + \cos(x) + x\,\mathrm{sen}(x) \right) \Bigg|_0^{\pi/2} \\
&= \frac{\pi^6}{768} + \frac{\pi}{2} - 1
\end{aligned}$$

□

3.2. Integral de Riemann en $\mathbb{R}^n$

Las demostraciones que no se incluyen en esta sección son idénticas a las hechas en la sección 3.1, acerca de la integral de Riemann en $\mathbb{R}^2$.

3.2.1. Definiciones

Definición 3.10. $R \subseteq \mathbb{R}^n$ es un rectángulo si y sólo si $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ con $a_i, b_i \in \mathbb{R}$.

Definición 3.11. Para un rectángulo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ se define su volumen como

$$V(R) = \prod_{i=1}^{n} b_i - a_i$$

Definición 3.12. Para $m \in \mathbb{N}$, la m-equipartición del rectángulo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ es $\{I_1 \times \cdots \times I_n : I_i \in P_i, i = 1, \ldots, n\}$ con P_i m-equipartición del intervalo $[a_i, b_i], i = 1, \ldots, n$. La denotaremos $P_m(R)$.

Dado un rectángulo $R \subseteq \mathbb{R}^n$ y $m \in \mathbb{N}$, decimos que $(\boldsymbol{c}_P)_{P \in P_m(R)}$ es una selección (para $P_m(R)$) si $(\forall P \in P_m(R)) \boldsymbol{c}_P \in P$.

Notar que cada elemento de una m-equipartición es un rectángulo y que una m-equipartición es finita, luego la siguiente definición tiene sentido:

Definición 3.13. Sea $m \in \mathbb{N}$, $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$ y una selección $(\boldsymbol{c}_P)_{P \in P_m(R)}$. Se define la suma de Riemann asociada a f y $(\boldsymbol{c}_P)_{P \in P_m(R)}$ como:

$$S\left(f, (\boldsymbol{c}_P)_{P \in P_m(R)}\right) = \sum_{P \in P_m(R)} f(\boldsymbol{c}_P) V(P)$$

Definición 3.14. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$. Decimos que f es Riemann integrable en R si y sólo si $(\exists S \in \mathbb{R})$

$$\left| S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right) - S\right| < \varepsilon$$

para toda elección de los $(\boldsymbol{c}_P)_{P\in P_m(R)}$.

S se llama integral (de Riemann) de f sobre R y se denota:

$$\int_R f$$

La integral de f también se anota:

$$\int_R f(\boldsymbol{x})\, d\boldsymbol{x}$$

3.2.2. Propiedades Básicas

Las propiedades que enunciaremos tienen una demostración análoga a las que ya vimos en el caso de $\mathbb{R}^2$.

Proposición 3.8. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$. Si f es integrable en R entonces f es acotada en R.

Proposición 3.9. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$. Son equivalentes:

1. $\forall\varepsilon > 0, \exists m_0 \in \mathbb{N}\ \forall k, m \geq m_0, \forall(\boldsymbol{c}'_P)_{P\in P_k(R)}$ selección, $\forall(\boldsymbol{c}_P)_{P\in P_m(R)}$ selección
$$\left| S\left(f, (\boldsymbol{c}'_P)_{P\in P_k(R)}\right) - S\left(f, (\boldsymbol{c}_P)_{P\in P_m(R)}\right)\right| < \varepsilon$$
2. f es integrable en R.

Proposición 3.10. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$. Si f es continua en R entonces f es integrable en R.

Proposición 3.11. Sean $R \subseteq \mathbb{R}^n$ rectángulo, $f, g : R \to \mathbb{R}$ funciones integrables y $c \in \mathbb{R}$. Se tienen las siguientes propiedades:

1. Linealidad: $f + cg$ es integrable, entonces
$$\int_R f + cg = \int_R f + c\int g$$
2. Monotonía: si $(\forall x \in R) f(x) \leq g(x)$, entonces
$$\int_R f \leq \int_R g$$
3. $|f|$ es integrable en R, entonces
$$\left|\int_R f\right| \leq \int_R |f|$$

Proposición 3.12. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$. Si f es integrable en R entonces

$$V(R) \inf_{\boldsymbol{x}\in R} f(\boldsymbol{x}) \leq \int_R f \leq V(R) \sup_{\boldsymbol{x}\in R} f(\boldsymbol{x})$$

3.2.3. Integración de sucesiones de funciones

Proposición 3.13. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $(f_k)_{k\in\mathbb{N}}$ sucesión de funciones, $f_k : R \to \mathbb{R}$, que convergen uniformemente en R a $f : R \to \mathbb{R}$. Entonces f es integrable en R y

$$\int_R f = \lim_{k\to\infty} \int_R f_k$$

3.2.4. Extensión de la clase de funciones integrables

Recordar que el grafo de una función $g : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es

$$\operatorname{grafo}(g) = \{(\boldsymbol{x}, g(\boldsymbol{x})) \in \mathbb{R}^{n+m} : \boldsymbol{x} \in A\}$$

La demostración del siguiente lema es análoga a la del lema 3.2.

Lema 3.3. *Sea $R' \subseteq \mathbb{R}^{N-1}$ rectángulo, $f : R' \to [a,b]$ continua. Denotemos $R = R' \times [a,b]$. Entonces*

$$\lim_{m\to\infty} \sum_{\substack{P\in P_m(R)\\ P\cap\operatorname{grafo}(g)\neq\varnothing}} V(P) = 0$$

Demostración. Sea $\varepsilon > 0$. Como f continua en R' compacto, se tiene que f es uniformemente continua en R'. Luego $(\exists\delta > 0)(\forall \boldsymbol{x}, \boldsymbol{y} \in R')$

$$\|\boldsymbol{x} - \boldsymbol{y}\| < \delta \Rightarrow \left| f(\boldsymbol{x}) - f(\boldsymbol{y}) \right| < \varepsilon$$

Así, escojamos $m_0 \in \mathbb{N}$ tal que $V(R)/m_0 \leq \varepsilon$ y $\operatorname{diam}(P) < \delta$, para cualquier $P \in P_m(R')$. Entonces, se cumple que $(\forall m \geq m_0)$

$$\begin{aligned}
\sum_{\substack{P\in P_m(R)\\ P\cap\operatorname{grafo}(g)\neq\varnothing}} V(P) &= \frac{V(R)}{m^N}\left|\{P \in P_m(R) : P \cap \operatorname{grafo}(g) \neq \varnothing\}\right| \\
&\leq \frac{V(R)}{m^N} m^{N-1}\left(\frac{\varepsilon}{(b-a)/m} + 1\right) \\
&= \varepsilon V(R') + \frac{V(R)}{m} \\
&\leq \varepsilon(V(R') + 1)
\end{aligned}$$

■

Proposición 3.14. Sea $R' \subseteq \mathbb{R}^{N-1}$ rectángulo, $R = R' \times [a,b] \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$ acotada en R y continua en $R \to \operatorname{grafo}(g)$, donde $g : R' \to [a,b]$ es una función continua. Entonces f es integrable en R.

Corolario 3.2. *Sea $R' \subseteq \mathbb{R}^{N-1}$ rectángulo, $R = R' \times [a,b] \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$ acotada en R y continua en $R \to \bigcup_{i=1}^{n} \operatorname{grafo}(g_i)$, donde $g_i : R' \to [a,b]$ es continua. Entonces f es integrable en R.*

3.2.5. Teorema de Fubini

Lema 3.4. *Sea $R \subseteq \mathbb{R}^n$ rectángulo, $f : R \to \mathbb{R}$, $m^* \in \mathbb{N}$. Si f es integrable en R entonces*

$$\sum_{S \in P_{m^*}(R)} V(S) \inf_{\boldsymbol{x} \in S} f(\boldsymbol{x}) \leq \int_R f \leq \sum_{S \in P_{m^*}(R)} V(S) \sup_{\boldsymbol{x} \in S} f(\boldsymbol{x})$$

Demostración. En efecto, es claro que $\forall \boldsymbol{x} \in R$

$$\sum_{S \in P_{m^*}(R)} \mathbf{1}_S(\boldsymbol{x}) \inf_{\boldsymbol{x} \in S} f(\boldsymbol{x}) \leq f(\boldsymbol{x}) \leq \sum_{S \in P_{m^*}(R)} \mathbf{1}_S(\boldsymbol{x}) \sup_{\boldsymbol{x} \in S} f(x).$$

Integrando se concluye la desigualdad buscada, notando que para $S \in P_{m^*}(R)$ se tiene

$$\int_R \mathbf{1}_S = \int_S 1 = V(S)$$

■

Teorema 3.2. (Teorema de Fubini en $\mathbb{R}^n$)

Sean $M, N \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ rectángulos, notemos $R = A \times B \subseteq \mathbb{R}^{n+m}$ rectángulo. Sea $f : R \to \mathbb{R}$.

1. *Si f integrable en R y $(\forall \boldsymbol{x} \in A) f(\boldsymbol{x}, \cdot)$ integrable en B, entonces*

$$\int_R f = \int_A \left(\int_B f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y} \right) d\boldsymbol{x}$$

2. *Si f continua en R, entonces*

$$\int_R f = \int_A \left(\int_B f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{y} \right) d\boldsymbol{x} = \int_B \left(\int_A f(\boldsymbol{x}, \boldsymbol{y})\, d\boldsymbol{x} \right) d\boldsymbol{y}$$

Demostración. La segunda parte es consecuencia directa de la primera, en virtud de la proposición 3.3. Probemos la primera parte.

Definamos la función $I : A \to \mathbb{R}$ dada por $I(x) = \int_B f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y}$. Luego basta probar que I es integrable en A y

$$\int_A I = \int_R f$$

En efecto, sea $\varepsilon > 0$ arbitrario. Por una parte, por definición de integrabilidad, $(\exists m_0 \in \mathbb{N})(\forall m \geq m_0)$

$$\left| \int_R f - \sum_{Q \in P_m(R)} f(\boldsymbol{c}_Q) V(Q) \right| < \varepsilon \tag{3.7}$$

para toda elección de los $(\boldsymbol{c}_Q)_{Q \in P_m(R)}$.

Por otra parte, sea $m \geq m_0$, $(\boldsymbol{c}_{Q_A})_{Q_A \in P_m(a)}$ selección arbitraria. Notar que

$$P_m(R) = \left\{ Q_A \times Q_B : Q_A \in P_m(a), Q_B \in P_m(B) \right\}$$

Consideremos la selección $(\boldsymbol{c}_Q^*)_{Q\in P_m(R)}$ dada por (para $Q_A \in P_m(a), Q_B \in P_m(B)$) $\boldsymbol{c}^*_{Q_A\times Q_B} = (\boldsymbol{c}_{Q_A}, y^*)$, con y^* tal que $f(\boldsymbol{c}^*_{Q_A\times Q_B}) \geq \sup_{y\in Q_B} f(\boldsymbol{c}_{Q_A}, y) - \varepsilon$. Así, en virtud del lema 3.4 se tiene que

$$\begin{aligned}
\sum_{Q_A\in P_m(a)} I(\boldsymbol{c}_{Q_A})V(Q_A) - \sum_{Q\in P_m(R)} f(\boldsymbol{c}_Q^*)V(Q)
&\leq \sum_{\substack{Q_A\in P_m(a)\\ Q_B\in P_m(B)}} \sup_{\boldsymbol{y}\in Q_B} f(\boldsymbol{c}_{Q_A}, \boldsymbol{y})V(Q_A)V(Q_B) \\
&\quad - \sum_{\substack{Q_A\in P_m(a)\\ Q_B\in P_m(B)}} f(c^*_{Q_A\times Q_B})V(Q_A)V(Q_B) \\
&= \sum_{\substack{Q_A\in P_m(a)\\ Q_B\in P_m(B)}} \big(\sup_{\boldsymbol{y}\in Q_B} f(\boldsymbol{c}_{Q_A}, \boldsymbol{y}) - f(\boldsymbol{c}^*_{Q_A\times Q_B})\big)V(Q_A)V(Q_B) \\
&\leq \varepsilon V(R)
\end{aligned}$$

Combinando con (3.7) se obtiene:

$$\sum_{Q_A\in P_m(a)} I(\boldsymbol{c}_{Q_A})V(Q_A) - \int_R f \leq \varepsilon\big(V(R)+1\big)$$

Escogiendo $(\boldsymbol{c}_Q^*)_{Q\in P_m(R)}$ tal que $f(\boldsymbol{c}^*_{Q_A\times Q_B}) \leq \inf_{\boldsymbol{y}\in Q_B} f(\boldsymbol{c}_{Q_A}, \boldsymbol{y}) + \varepsilon$, un cálculo análogo permite obtener:

$$-\varepsilon\big(V(R)+1\big) \leq \sum_{Q_A\in P_m(a)} I(\boldsymbol{c}_{Q_A})V(Q_A) - \int_R f$$

En conclusión,

$$\left|\sum_{Q_A\in P_m(a)} I(\boldsymbol{c}_{Q_A})V(Q_A) - \int_R f\right| \leq \varepsilon\big(V(R)+1\big)$$

$\blacksquare$

Ejemplo 3.6. Calcular $\int_B f$ con $B = [0,1]^3 \times [0,10]$ y $f(x,y,z,t) = t(x^2+y^2+z^2)$.

Solución. En virtud del teorema de Fubini

$$\begin{aligned}
\int_B f &= \int_0^{10}\int_0^1\int_0^1\int_0^1 t(x^2+y^2+z^2)\,dx\,dy\,dz\,dt \\
&= \int_0^{10}\int_0^1\int_0^1 \left(t\frac{x^3}{3} + ty^2x + tz^2x\right)\Big|_{x=0}^{1} dy\,dz\,dt \\
&= \int_0^{10}\int_0^1\int_0^1 \left(\frac{t}{3} + ty^2 + tz^2\right) dy\,dz\,dt \\
&= \int_0^{10}\int_0^1 \left(\frac{t}{3}y + t\frac{y^3}{3} + tz^2y\right)\Big|_{y=0}^{1} dz\,dt \\
&= \int_0^{10}\int_0^1 \left(\frac{t}{3} + \frac{t}{3} + tz^2\right) dz\,dt \\
&= \int_0^{10} t\,dt \\
&= 50.
\end{aligned}$$

□

3.2.6. Integral en $\mathbb{R}^n$ sobre dominios generales

Definición 3.15. Sea $R \subseteq \mathbb{R}^n$ rectángulo, $A \subseteq R$, $f : A \to \mathbb{R}$. Denotemos $f_0 : R \to \mathbb{R}$ a la función dada por, para $x \in R$:

$$f_0(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) & \text{si } \boldsymbol{x} \in A \\ 0 & \text{si } \boldsymbol{x} \notin A \end{cases}$$

Si f_0 es integrable sobre R entonces se dice que f es integrable sobre A y:

$$\int_A f = \int_R f_0$$

Definición 3.16. La función indicatriz de un conjunto $A \subseteq \mathbb{R}^n$ es análoga al caso de $\mathbb{R}^2$ y se define:

$$\mathbf{1}_A(\boldsymbol{x}) = \begin{cases} 1 & \text{si } \boldsymbol{x} \in A \\ 0 & \text{si } \boldsymbol{x} \notin A \end{cases}$$

Definición 3.17. Decimos que $D \subseteq \mathbb{R}^3$ es

1. Dominio de tipo 1 si y sólo si alguna de las siguientes afirmaciones es cierta:

   a) $\exists a, b \in \mathbb{R}$, $\exists \phi_1, \phi_2 : [a,b] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

   $$\begin{aligned} D = \{(x,y,z) \in \mathbb{R}^3 : a \le x \le b,& \\ \phi_1(x) \le y \le \phi_2(x),& \\ \gamma_1(x,y) \le z \le \gamma_2(x,y)\}& \end{aligned}$$

   donde $D' = \{(x,y) \in \mathbb{R}^2 : a \le x \le b, \phi_1(x) \le y \le \phi_2(x)\}$

b) $\exists c, d \in \mathbb{R}$, $\exists \psi_1, \psi_2 : [c, d] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

$$\begin{aligned} D = \{(x, y, z) \in \mathbb{R}^3 : c \leq y \leq d,\\ \psi_1(y) \leq x \leq \psi_2(y),\\ \gamma_1(x, y) \leq z \leq \gamma_2(x, y)\} \end{aligned}$$

donde $D' = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$.

2. Dominio de tipo 2 si y sólo si alguna de las siguientes afirmaciones es cierta:

a) $\exists a, b \in \mathbb{R}$, $\exists \phi_1, \phi_2 : [a, b] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

$$\begin{aligned} D = \{(x, y, z) \in \mathbb{R}^3 : a \leq x \leq b,\\ \phi_1(x) \leq z \leq \phi_2(x),\\ \gamma_1(x, z) \leq y \leq \gamma_2(x, z)\} \end{aligned}$$

donde $D' = \{(x, z) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq z \leq \phi_2(x)\}$.

b) $\exists c, d \in \mathbb{R}$, $\exists \psi_1, \psi_2 : [c, d] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

$$\begin{aligned} D = \{(x, y, z) \in \mathbb{R}^3 : c \leq z \leq d,\\ \psi_1(z) \leq x \leq \psi_2(z),\\ \gamma_1(x, z) \leq y \leq \gamma_2(x, z)\} \end{aligned}$$

donde $D' = \{(x, z) \in \mathbb{R}^2 : c \leq z \leq d, \psi_1(z) \leq x \leq \psi_2(z)\}$.

3. Dominio de tipo 3 si y sólo si alguna de las siguientes afirmaciones es cierta:

a) $\exists a, b \in \mathbb{R}$, $\exists \phi_1, \phi_2 : [a, b] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

$$\begin{aligned} D = \{(x, y, z) \in \mathbb{R}^3 : a \leq y \leq b,\\ \phi_1(y) \leq z \leq \phi_2(y),\\ \gamma_1(y, z) \leq x \leq \gamma_2(y, z)\} \end{aligned}$$

donde $D' = \{(y, z) \in \mathbb{R}^2 : a \leq y \leq b, \phi_1(y) \leq z \leq \phi_2(y)\}$.

b) $\exists c, d \in \mathbb{R}$, $\exists \psi_1, \psi_2 : [c, d] \to \mathbb{R}$ funciones continuas, $\exists \gamma_1, \gamma_2 : D' \to \mathbb{R}$ funciones continuas tales que

$$\begin{aligned} D = \{(x, y, z) \in \mathbb{R}^3 : c \leq z \leq d,\\ \psi_1(z) \leq y \leq \psi_2(z),\\ \gamma_1(y, z) \leq x \leq \gamma_2(y, z)\} \end{aligned}$$

donde $D' = \{(y, z) \in \mathbb{R}^2 : c \leq z \leq d, \psi_1(z) \leq y \leq \psi_2(z)\}$.

4. Dominio de tipo 4 si y sólo si D es de tipo 1, 2 y 3.

5. Dominio elemental si y sólo si D es de tipo 1, 2 ó 3.

Si $R \subseteq \mathbb{R}^n$ es un rectángulo y $D \subseteq R$ es una región elemental entonces f_0 (dada por la definición 3.15) es continua en R salvo sobre la unión finita de grafos de funciones. Luego $\int_D f$ existe.

Ejemplo 3.7. Sea $W \subseteq \mathbb{R}^3$ la región comprendida entre los planos $x = 0$, $y = 0$, $z = 2$ y la superficie $z = x^2 + y^2$. Calcular $\int_W x$.

Solución. ¿Cómo describir el conjunto W? Podemos describirlo como un dominio de tipo 1, es decir:

$$W = \{(x, y, z) \in \mathbb{R}^3 : 0 \leq x \leq \sqrt{2}, 0 \leq y \leq \sqrt{2 - x^2}, x^2 + y^2 \leq z \leq 2\}$$

Luego, gracias al teorema de Fubini (teorema 3.2) se tiene:

$$\begin{aligned}
\int_W x &= \int_0^{\sqrt{2}} \int_0^{\sqrt{2-x^2}} \int_{x^2+y^2}^{2} x \, dz \, dy \, dx \\
&= \int_0^{\sqrt{2}} \int_0^{\sqrt{2-x^2}} 2x - x(x^2 + y^2) \, dy \, dx \\
&= \int_0^{\sqrt{2}} \left(2xy - x^3y - x\frac{y^3}{3} \right) \Bigg|_{y=0}^{\sqrt{2-x^2}} dx \\
&= \int_0^{\sqrt{2}} x \left(2\sqrt{2 - x^2} - x^2\sqrt{2 - x^2} - \frac{1}{3}(2 - x^2)^{3/2} \right) dx \\
&= \int_0^{\sqrt{2}} x \left(\frac{2}{3}(2 - x^2)^{3/2} \right) dx
\end{aligned}$$

y, haciendo el cambio de variable $u = 2 - x^2$:

$$\begin{aligned}
&= \int_0^2 \frac{2}{3}\frac{1}{2} u^{3/2} du \\
&= \frac{1}{3}\frac{2}{5} u^{5/2} \Big|_0^2 \\
&= \frac{2}{15} 2^{5/2} \\
&= \frac{8\sqrt{2}}{15}
\end{aligned}$$

□

3.3. Reglas de derivación adicionales

Teorema 3.3. (Regla de Leibniz de derivación)
Sea $f : \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathcal{C}^1$. $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ funciones diferenciables. Entonces la función $F : \mathbb{R} \to \mathbb{R}$ definida por

$$F(x) = \int_{\alpha(x)}^{\beta(x)} f(x, t) dt$$

es diferenciable y su derivada es

$$\frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(x, t) dt = f(x, \beta(x))\beta'(x) - f(x, \alpha(x))\alpha'(x) + \int_{\alpha(x)}^{\beta(x)} \frac{\partial}{\partial x} f(x, t) dt$$

Demostración. Definamos la función

$$G(z,x)=\int_0^z f(x,t)dt$$

por el 1^{er} teorema fundamental del cálculo (teorema 1.8) sabemos que

$$\frac{\partial}{\partial z}G(z,x)=f(x,z)$$

demostremos que

$$\frac{\partial}{\partial x}G(z,x)=\int_0^z \frac{\partial}{\partial x}f(x,t)dt$$

para ello notemos que

$$\begin{aligned}
& G(z,x+h)-G(z,x)-h\int_0^z \frac{\partial}{\partial x}f(x,t)dt \\
= & \int_0^z f(x+h,t)+f(x,t)-h\frac{\partial}{\partial x}f(x,t)dt \\
= & \int_0^z\int_0^1 (\frac{\partial}{\partial x}f(x+yh,t)-\frac{\partial}{\partial x}f(x,t))hdydt
\end{aligned}$$

La función $\partial f/\partial x(\cdot,\cdot)$ es continua, y por lo tanto, uniformemente continua sobre $[x-|h|,x+|h|]\times[0,z]$, asi, dado $\varepsilon>0$, existe $\delta>0$ tal que si $\|(x,y)-(\bar{x},\bar{y})\|<\delta$ entonces $|\partial f/\partial x(x,y)-\partial f/\partial x(\bar{x},\bar{y})|<\varepsilon/z$. Entonces si $|h|<\delta$ se tiene que

$$|\frac{\partial}{\partial x}f(x+yh,t)-\frac{\partial}{\partial x}f(x,t)|<\varepsilon\ \ \forall t\in[0,z]\ \forall y\in[0,1]$$

lo que implica que

$$|G(z,x+h)-G(z,x)-h\int_0^z \frac{\partial}{\partial x}f(x,t)dt|<\varepsilon|h|$$

si $|h|<\delta$ de donde se sigue el resultado. Luego

$$\int_{\alpha(x)}^{\beta(x)} f(x,t)dt=G(\beta(x),x)-G(\alpha(x),x)$$

y aplicando el resultado anterior se tiene que

$$\begin{aligned}
& \frac{d}{dx}\int_{\alpha(x)}^{\beta(x)} f(x,t)dt \\
= & \frac{\partial}{\partial z}G(\beta(x),x)\beta'(x)+\frac{\partial}{\partial x}G(\beta(x),x)\frac{dx}{dx}-\frac{\partial}{\partial z}G(\alpha(x),x)\alpha'(x)-... \\
& ...-\frac{\partial}{\partial x}G(\beta(x),x)\frac{dx}{dx} \\
= & f(x,\beta(x))\beta'(x)-f(\alpha(x),x)\alpha'(x)+\int_0^{\beta(x)}\frac{\partial f}{\partial x}(x,t)dt-\int_0^{\alpha(x)}\frac{\partial f}{\partial x}(x,t)dt \\
= & f(x,\beta(x))\beta'(x)-f(x,\alpha(x))\alpha'(x)+\int_{\alpha(x)}^{\beta(x)}\frac{\partial}{\partial x}f(x,t)dt
\end{aligned}$$

■

Teorema 3.4. *Sean $A \subseteq \mathbb{R}^n$ y $B \subseteq \mathbb{R}^m$ abiertos y $f : A \times B \to \mathbb{R}$ de clase $\mathcal{C}^1$. Sea $Q \subset B$ conjunto elemental cerrado. Entonces la función*

$$F(\boldsymbol{x}) = \int_Q f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y}$$

es de clase $\mathcal{C}^1$ y

$$\frac{\partial}{\partial x_i} F(\boldsymbol{x}) = \int_Q \frac{\partial}{\partial x_i} f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y}$$

para todo $i \in \{1, ..., n\}$

Demostración. Sea $\boldsymbol{x}_0 \in A$

$$\begin{aligned}
& F(\boldsymbol{x}_0 + \boldsymbol{h}) - F(\boldsymbol{x}_0) - \sum_{i=1}^n h_i \int_Q \frac{\partial}{\partial x_i} f(\boldsymbol{x}_0, \boldsymbol{y}) d\boldsymbol{y} \\
= & \int_Q f(\boldsymbol{x}_0 + \boldsymbol{h}, \boldsymbol{y}) - f(\boldsymbol{x}_0, \boldsymbol{y}) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0, \boldsymbol{y}) d\boldsymbol{y} \\
= & \int_Q (f(\boldsymbol{x}_0 + \boldsymbol{h}, \boldsymbol{y}) - f(\boldsymbol{x}_0, \boldsymbol{y}) - \nabla_x f(\boldsymbol{x}_0, \boldsymbol{y}) \cdot \boldsymbol{h}) d\boldsymbol{y} \\
= & \int_Q \int_0^1 \left[(\nabla_x f(\boldsymbol{x}_0 + t\boldsymbol{h}, \boldsymbol{y}) - \nabla_x f(\boldsymbol{x}_0, \boldsymbol{y})) \cdot \boldsymbol{h} \right] dt d\boldsymbol{y}
\end{aligned}$$

La función $\nabla_x f(\cdot, \cdot)$ es continua en $\bar{B}(\boldsymbol{x}_0, 1) \times Q$ que es compacto, por cual que es uniformemente continua. Luego dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $\|(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{x}^*, \boldsymbol{y}^*)\| < \delta \Rightarrow \|\nabla_x f(\boldsymbol{x}, \boldsymbol{y}) - \nabla_x f(\boldsymbol{x}^*, \boldsymbol{y})\| < \varepsilon/\text{vol}(Q)$. entonces si $\|\boldsymbol{h}\| < \delta$ se tiene que

$$\|\nabla_x f(\boldsymbol{x}_0 + t\boldsymbol{h}, \boldsymbol{y}) - \nabla_x f(\boldsymbol{x}_0, \boldsymbol{y})\| < \frac{\varepsilon}{\text{vol}(Q)} \; \forall\, t \in [0, 1] \, \forall\, \boldsymbol{y} \in Q$$

lo que implica que

$$\begin{aligned}
& |F(\boldsymbol{x}_0 + \boldsymbol{h}) - F(\boldsymbol{x}_0) - \sum_{i=1}^n h_i \int_Q \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0, \boldsymbol{y}) d\boldsymbol{y}| \\
\leq & \int_Q \int_0^1 \|\nabla f(\boldsymbol{x}_0 + t\boldsymbol{h}, \boldsymbol{y}) - \nabla f(\boldsymbol{x}_0, \boldsymbol{y})\| \|\boldsymbol{h}\| dt d\boldsymbol{y} \\
\leq & \int_Q \int_0^1 \frac{\varepsilon}{\text{vol}(Q)} \|\boldsymbol{h}\| = \varepsilon \|\boldsymbol{h}\|
\end{aligned}$$

de este modo

$$\left| F(\boldsymbol{x}_0 + \boldsymbol{h}) - F(\boldsymbol{x}_0) - \sum_{i=1}^n h_i \int_Q \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0, \boldsymbol{y}) d\boldsymbol{y} \right| \leq \varepsilon \|\boldsymbol{h}\|$$

siempre que $\|\boldsymbol{h}\| < \delta$, es decir, F es diferenciable es $\boldsymbol{x}_0$ y

$$\frac{\partial}{\partial x_i} F(\boldsymbol{x}_0) = \int_Q \frac{\partial}{\partial x_i} f(\boldsymbol{x}_0, \boldsymbol{y}) d\boldsymbol{y}$$

para todo $i \in \{1, ..., n\}$. La continuidad de las derivadas parciales queda de *ejercicio* . ■

3.4. Teorema del cambio de variable

Recordemos que en el caso de una variable se tiene que si $\sigma : [a,b] \to [c,d]$ es biyectiva (más algunas hipótesis adicionales) entonces

$$\int_c^d f(t)\,dt = \int_{\sigma^{-1}(c)}^{\sigma^{-1}(d)} f(\sigma(s))\sigma'(s)\,ds$$

Equivalentemente podemos escribir:

$$\int_{\sigma([a,b])} f(t)\,dt = \int_{[a,b]} f(\sigma(s))\,|\sigma'(s)|\;ds.$$

Nuestro propósito es extender esta fórmula a varias variables. Vamos a comenzar con la transformación más simple, la lineal. Sea $T : D = [0,1]^2 \to D^* = T(D)$, $T(x,y) = A\binom{x}{y}$ con $A \in M_{2,2}(\mathbb{R})$

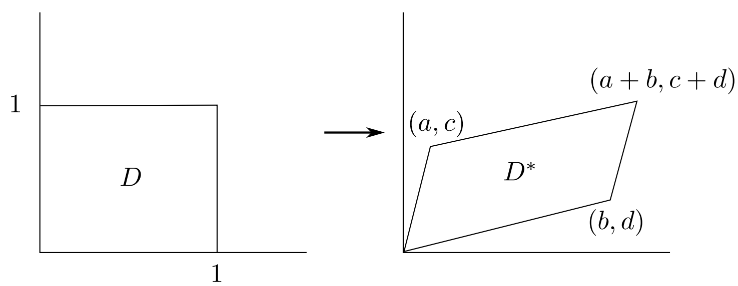


Figura 3.8: Cambio de variable, caso de una transformación lineal

Se tiene que $V(D) = 1$ y

$$\begin{aligned} V(D^*) &= \left| \frac{(b,d)\cdot(-c,a)}{\sqrt{a^2+c^2}} \sqrt{a^2+c^2} \right| \\ &= |ad - bc| \\ &= |\det(A)| \end{aligned}$$

Así

$$V(D^*) = V(D)\,|\det|$$

Es fácil ver que una fórmula análoga se cumple para cualquier rectángulo en $\mathbb{R}^n$. Más aún, para $f : D^* \to \mathbb{R}$, de acuerdo a la definición de integral de Riemann es razonable escribir lo siguiente:

$$\begin{aligned} \int_{D^*} f &\approx \lim_m \sum_{P \in P_m(D)} f(T(\boldsymbol{c}_P))V(T(P)) \qquad (\boldsymbol{c}_P \in P) \\ &\approx \lim_m \sum_{P \in P_m(D)} f(T(\boldsymbol{c}_P))V(P)\,|\det(A)| \\ &\approx \int_D f \circ T\,|\det(A)| \end{aligned}$$

Para una transformación más general (no lineal), se puede pensar que la diferencial nos da una aproximación lineal de la transformación en torno a un punto, luego es razonable pensar que el jacobiano de la transformación desempeñará el papel de la matriz A en la fórmula más general. En efecto, se tiene el siguiente teorema (para los detalles, véase la sección 6.2):

Teorema 3.5. (Teorema del cambio de variables)
Sea Ω v-medible y $f : U \to V$ un difeomorfismo, $\Omega \subseteq U$. Sea $g : f(\Omega) \to \mathbb{R}$ v-integrable. Entonces $g \circ f : \Omega \to \mathbb{R}$ es v-integrable y

$$\int_{f(\Omega)} g(x)dx = \int_{\Omega} g(f(y))|\det Df(y)|dy$$

Demostración. Por razones pedagógicas se hará posteriormente (teorema 6.3). ■

Ejemplo 3.8. (integral en coordenadas polares)
Calcular

$$\int_C \log(x^2 + y^2)\, dx\, dy$$

donde

$$C = \{(x, y) : a^2 \leq x^2 + y^2 \leq b^2, x \geq 0, y \geq 0\}$$

Solución. Vamos a considerar el cambio de variable

$$\begin{aligned} x &= r\cos(\theta) \\ y &= r\,\mathrm{sen}(\theta) \end{aligned}$$

es decir, la transformación $T(r, \theta) = (r\cos(\theta), r\,\mathrm{sen}(\theta))$. Notar que:

$$D(r, \theta) = \begin{bmatrix} \cos(\theta) & -r\,\mathrm{sen}(\theta) \\ \mathrm{sen}(\theta) & r\cos(\theta) \end{bmatrix}$$

y así:

$$\det DT(r, \theta) = r$$

Denotemos

$$\begin{aligned} D &= T^{-1}(C) \\ &= \{(r, \theta) \in \mathbb{R}^2 : a \leq r \leq b, 0 \leq \theta \leq \pi/2\} \\ &= [a, b] \times [0, \pi/2] \end{aligned}$$

Luego, en virtud del teorema del cambio de variable :

$$\int_{T(D)} \log(x^2 + y^2)\, dx\, dy = \int_D \log(r^2) r\, dr\, d\theta$$

y, con Fubini:

$$\begin{aligned} &= \int_a^b \int_0^{\pi/2} \log(r^2) r\, dr\, d\theta \\ &= \frac{\pi}{2} \frac{1}{2} \int_{a^2}^{b^2} \log s\, ds \\ &= \frac{\pi}{4}(b^2 \log b^2 - b^2 - a^2 \log a^2 + a^2) \end{aligned}$$

□

3.5. Aplicaciones

3.5.1. Centro de masa

Sea $D \subseteq \mathbb{R}^2$ una placa y $\rho : D \to \mathbb{R}$ la densidad (de masa). Entonces la masa total de la placa está dada por:

$$\iint_D \rho$$

y las coordenadas $(\bar{x}, \bar{y})$ del centro de masa de la placa están dadas por:

$$\bar{x} = \frac{\iint_D x\rho(x,y)\,dy\,dx}{M}$$
$$\bar{y} = \frac{\iint_D y\rho(x,y)\,dy\,dx}{M}$$

El caso tridimensional es análogo.

Ejemplo 3.9. Hallar el centro de masa de una lámina triangular con vértices $(0,0)$, $(1,0)$ y $(0,2)$ si la función de densidad es $\rho(x,y) = 1 + 3x + y$.

Solución. A partir de los vértices la hipotenusa del tríangulo queda descrita por la ecuación $y = 2 - 2x$. Luego la masa total corresponde a

$$\begin{aligned} M &= \iint_D \rho(x,y) = \int_0^1 \int_0^{2-2x} (1+3x+y)dydx = \int_0^1 \left[y + 3xy + \frac{y^2}{2} \right]_{y=0}^{y=2-2x} dx \\ &= 4\int_0^1 (1-x^2)dx = 4\left[x - \frac{x^3}{3} \right]_0^1 = \frac{8}{3} \end{aligned}$$

Luego el centro de masa está dado por las ecuaciones

$$\begin{aligned} \bar{x} &= \frac{1}{M}\iint_D x\rho(x,y)dydx = \frac{3}{8}\int_0^1 \int_0^{2-2x} (x + 3x^2 + xy)dydx \\ &= \frac{3}{8}\int_0^1 \left[y + 3x^2y + \frac{xy^2}{2} \right]_{y=0}^{y=2-2x} dx = \frac{3}{2}\int_0^1 (x - x^3)dx = \left[\frac{x^2}{2} - \frac{x^4}{4} \right]_0^1 = \frac{3}{8} \end{aligned}$$

$$\begin{aligned} \bar{y} &= \frac{1}{M}\iint_D y\rho(x,y)dydx = \frac{3}{8}\int_0^1 \int_0^{2-2x} (y + 3xy + y^2)dydx \\ &= \frac{3}{8}\int_0^1 \left[\frac{y^2}{2} + \frac{3xy^2}{2} + \frac{y^3}{3} \right]_{y=0}^{y=2-2x} dx = \frac{1}{4}\int_0^1 (7 - 9x - 3x^2 + 5x^3)dx \\ &= \frac{1}{4}\left[7x - \frac{9x^2}{2} - x^3 + \frac{5x^4}{4} \right]_0^1 = \frac{11}{16} \end{aligned}$$

por lo tanto, el centro de masa se ubica en las coordenadas $(\bar{x}, \bar{y}) = (\frac{3}{8}, \frac{11}{16})$ $\square$

3.5.2. Momento de inercia

Sea $W \subseteq \mathbb{R}^3$ un sólido de densidad $\rho : W \to \mathbb{R}$. Entonces los momentos de inercia están dados por:

$$I_x = \iiint_W \rho(x,y,z)(y^2+z^2)\,dz\,dy\,dx$$
$$I_y = \iiint_W \rho(x,y,z)(x^2+z^2)\,dz\,dy\,dx$$
$$I_z = \iiint_W \rho(x,y,z)(x^2+y^2)\,dz\,dy\,dx$$

Ejemplo 3.10. Hallar el centro de masa de la región W comprendida entre el plano xy y la semiesfera $x^2+y^2+z^2=1$, $z \geq 0$, de densidad $\rho(x,y)=1$, $\forall (x,y) \in W$.

Solución. En primer lugar, por simetría $\bar{x} = \bar{y} = 0$.

Para calcular $\bar{z}$, vamos a hacer un cambio de variable a coordenadas esféricas:

$$\begin{aligned} x &= r\operatorname{sen}(\phi)\cos(\theta) \\ y &= r\operatorname{sen}(\phi)\operatorname{sen}(\theta) \\ z &= r\cos(\phi) \end{aligned}$$

Así

$$\int_W z = \int_{T^{-1}(W)} r\cos(\phi)\,|\det DT|$$

$$DT(r,\phi,\theta) = \begin{bmatrix} \operatorname{sen}(\phi)\cos(\theta) & r\cos(\phi)\cos(\theta) & -r\operatorname{sen}(\phi)\operatorname{sen}(\theta) \\ \operatorname{sen}(\phi)\operatorname{sen}(\theta) & r\cos(\phi)\operatorname{sen}(\theta) & r\operatorname{sen}(\phi)\cos(\theta) \\ \cos(\phi) & -r\operatorname{sen}(\phi) & 0 \end{bmatrix}$$

$$\begin{aligned} \det DT(r,\phi,\theta) &= (-1)^{3+1}(-r\operatorname{sen}(\phi)\operatorname{sen}(\theta))(-r\operatorname{sen}^2(\phi)\operatorname{sen}(\theta) - r\cos^2(\phi)\operatorname{sen}(\theta)) \\ &\quad + (-1)^{3+2}(r\operatorname{sen}(\phi)\cos(\theta))(-r\operatorname{sen}^2(\phi)\cos(\theta) - r\cos^2(\phi)\cos(\theta)) \\ &= r^2\operatorname{sen}(\phi)\operatorname{sen}^2(\theta) + r^2\operatorname{sen}(\phi)\cos^2(\theta) \\ &= r^2\operatorname{sen}(\phi) \end{aligned}$$

Combinando lo anterior:

$$\begin{aligned} \int_W z &= \int_0^1 \int_0^{2\pi} \int_0^{\pi/4} (r\cos(\phi))r^2\operatorname{sen}(\phi)\,d\phi\,d\theta\,dr \\ &= \frac{r^4}{4}\bigg|_0^1 2\pi \int_0^{\pi/4} \operatorname{sen}(\phi)\cos(\phi)\,d\phi \\ &= \frac{\pi}{2}\frac{\operatorname{sen}^2(\phi)}{2}\bigg|_0^{\pi/4} \\ &= \frac{\pi}{8} \end{aligned}$$

Luego

$$\bar{z} = \frac{\frac{\pi}{8}}{\frac{4}{3}\frac{\pi}{2}} = \frac{3}{16}$$

$\square$

3.6. Comentarios acerca del capítulo

3.6.1. Extensión de la integral de Riemann

Algunos de los problemas que presenta la integral de Riemann son que, para ciertas aplicaciones, no integra una familia suficientemente amplia de funciones y que los teoremas de convergencia poseen hipótesis muy fuertes (convergencia uniforme en el caso de la proposición 3.13). Una construcción diferente, basada en la teoría de la medida, la constituye la integral de Lebesgue , que integra una familia mucho más amplia de funciones y cuenta con un teorema de convergencia basado en la convergencia puntual.

3.7. Ejercicios

Integral de Riemann

Ejercicio 1. Sea $R \subseteq \mathbb{R}^N$ rectángulo tal que $V(R) > 0$ y suponga que $f : R \to \mathbb{R}$ es continua en R. Suponga que para toda función continua $g : R \to \mathbb{R}$ se tiene

$$\int_R fg = 0$$

Pruebe que $f = 0$ en R.

Ejercicio 2. Sea el rectángulo $R = [0,1]^2 \subseteq \mathbb{R}^2$ y la función $f : R \to \mathbb{R}$ dada por:

$$f(x,y) = \begin{cases} 1 & \text{si } x \text{ es irracional} \\ 4y^3 & \text{si } x \text{ es racional} \end{cases}$$

1. Muestre que $\int_0^1 \int_0^1 f(x,y)\, dy\, dx$ existe y vale 1.
2. Muestre que $\int_R f$ no existe.

Ejercicio 3. Considere el rectángulo $D = [-1,1] \times [0,2]$ y R_n la partición uniforme con n^2 subrectangulos, es decir, todos los rectangulos de R_n tienen las mismas dimensiones. Dado esto, calcule las sumas superior e inferior $S(f,R_n), I(f,R_n)$ donde $f(x,y) = e^x y$.

Ejercicio 4. Suponga que $F \subseteq \mathbb{R}^N$ para $N \geq 2$ y que $f : R \to \mathbb{R}$ es Riemann integrable en F. Muestre que f puede ser no acotada en F.

Ejercicio 5. Pruebe que la integral de Riemann de una función en $\mathbb{R}^N$ es única.

Ejercicio 6. Justifique cuales de las siguiente funciones son integrables en $[0,1] \times [0,1]$:

1. $f(x,y) = \begin{cases} 0 & \text{si } (x,y) \notin A \\ 1 & \text{si } (x,y) \in A \end{cases} \quad , A = \{(x,y) : y \leq x^2\}$
2. $f(x,y) = \begin{cases} \frac{\text{sen}(x-y)}{x-y} & \text{si } x \neq y \\ 1 & \text{si } x = y \end{cases}$

Ejercicio 7. Sea $R = [0,1] \times [0;1]$ y $f : R \to \mathbb{R}$ definida por

$$f = \begin{cases} 0 & \text{si } x \leq y \\ 2 & \text{si } x > y \end{cases}$$

Demuestre que f es integrable y, usando sumas de Riemann, que $\int_R f = 1$

Ejercicio 8. Considere la siguiente integral

$$\int_{-5}^{2} \int_{0}^{\sqrt{25-x^2}} f(x,y)dydx + \int_{-2}^{0} \int_{2-\frac{x^2}{2}}^{\sqrt{25-x^2}} f(x,y)dydx + \int_{0}^{4} \int_{\frac{x}{4}+2}^{\sqrt{25-x^2}} f(x,y)dydx$$

Invierta el orden de integración usando Fubini.

Aplicaciones

Ejercicio 9. Calcule usando Fubini

$$\int_0^{\pi^2} \int_{\sqrt{y}}^{\pi} \frac{\operatorname{sen} x^2}{x^2} \sqrt[3]{y^2} dx dy$$

Ejercicio 10. Calcule el volumen del elipsoide dado por

$$E = \{(x, y, z) \in \mathbb{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\}$$

Indicación. Use coordenadas esféricas y cambio de variables.

Ejercicio 11. Calcule el volumen encerrado por el cono parabólico $x^2 + y^2 = z^2$ y por la bola $B(0, r)$ con $r > 0$.

Ejercicio 12. Evalúe la integral

$$\iint_{4x^2 - 8x + y^2 \leq 0} \sqrt{4x^2 + y^2} dx dy$$

Ejercicio 13. Calcule

$$\int_0^1 \int_{\sqrt{x}}^1 \frac{x}{\sqrt{x^2 + y^2}} dx dy$$

Ejercicio 14. Calcule el volumen del sólido que está limitado por las superficies

$$z - \sqrt{x^2 + y^2} = 0 \qquad \text{y} \qquad 2z - y - 2 = 0$$

Ejercicio 15. Hallar el volumen de la región acotada por los paraboloides

$$z = x^2 + y^2 \qquad \text{y} \qquad z = 10 - x^2 - 2y^2$$

Ejercicio 16. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por:

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

Pruebe que

$$\int_0^1 \int_0^1 f(x, y) dx dy \neq \int_0^1 \int_0^1 f(x, y) dy dx$$

¿Qué puede decir de este resultado en base al teorema de Fubini?

Ejercicio 17. Sea S la región del primer octante limitada por el plano $x + 2y + z = 2$ que se encuentra entre los planos $z = 0$ y $z = 1$.

1. Escriba las integrales iteradas que permiten calcular el volumen de la región.
2. Calcule el volumen que define la integral.

Ejercicio 18. Encuentre el área de la región definida por la curva

$$f(x) = (\operatorname{sen}(x)\cos(x), -\cos^2(x) + \operatorname{sen}^2(x))$$

en $[0, \pi]$.

Ejercicio 19. Calcular

$$\iiint\limits_S \sqrt{1 - \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)}$$

Donde $S = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$

Ejercicio 20. Calcular

$$\iiint\limits_S \exp -(11x^2 + 9y^2 + 15z^2 - 4xy - 20xz + 10yz)^2 dxdydz$$

Donde $S = 11x^2 + 9y^2 + 15z^2 - 4xy - 20xz + 10yz \leq 100$

Ejercicio 21. Calcular

$$\iiint\limits_S \frac{\exp -(11x^2 + 9y^2 + 15z^2 - 4xy - 20xz + 10yz)}{\sqrt{11x^2 + 9y^2 + 15z^2 - 4xy - 20xz + 10yz}} dxdydz$$

Donde $S = 11x^2 + 9y^2 + 15z^2 - 4xy - 20xz + 10yz \leq 80$

Ejercicio 22. Se define la masa total de un solido, con densidad $\lambda(x, y, z)$, como:

$$M = \iiint\limits_V \lambda(x, y, z) dxdydz$$

Además se definen las siguientes expresiones:

$$\begin{aligned} M_x &= \iiint\limits_V x\lambda(x, y, z) dxdydz \\ M_y &= \iiint\limits_V y\lambda(x, y, z) dxdydz \\ M_z &= \iiint\limits_V z\lambda(x, y, z) dxdydz \end{aligned}$$

El centro de masas corresponde a:

$$C_m = \frac{1}{M}(M_x, M_y, M_z)$$

Dado lo anterior, se pide que:

1. Considerando la función de densidad $\lambda(x,y,z) = \sqrt{\frac{x^2+y^2}{x^2+y^2+z^2}}$, calcule la masa total del cuerpo definido por
$$V = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 \leq 4, z \geq 0\}$$
2. Calcule el centro de masas para este mismo cuerpo de la parte 1.

Ejercicio 23. Calcular el volumen en $\mathbb{R}^5$ de
$$F = \{(x_1,x_2,x_3,x_4,x_5) : x_1^2+x_2^2+x_3^2 \leq 1, x_4^2+x_5^2 \leq 1\}$$

Ejercicio 24. Sea $f : A \subseteq \mathbb{R}^3 \to \mathbb{R}$ integrable y considere la integral
$$\int_{-1}^{1}\int_{0}^{x^2}\int_{0}^{y} f(x,y,z)dxdydz$$
Escriba la integral recién descrita en términos de integrales triples iteradas de la forma
$$\iiint f(x,y,z)dxdzdy \qquad \iiint f(x,y,z)dydzdx$$

Ejercicio 25. Calcular
$$\iint_S (1+xy)dxdy$$
donde $S = S_1 \cup S_2$ definidos por
$$S_1 = \begin{cases} -1 \leq x \leq 0 \\ 0 \leq y \leq (x+1)^2 \end{cases} \qquad S_2 = \begin{cases} 0 \leq x \leq 1 \\ 0 \leq y \leq (x-1)^2 \end{cases}$$

Reglas de derivación adicionales

Ejercicio 26. Sea $f(x,y,z,t) = \int_{e^{x+y+z}}^{\ell(1+|x+y|)} (xt + log_z(t) + z)dt$. Calcular $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial t}$ en aquellos puntos donde existan.

Ejercicio 27. Considere la siguiente ecuación integral
$$u(x) = 5 + \int_0^x \operatorname{sen}(u(s)+x)ds$$
En el capítulo 5 se demostró que tiene una única solución en $C([0,1/2],\mathbb{R})$.
Derivando dos veces la ecuación integral (justifique por que se puede derivar), encuentre la ecuación diferencial que satisface su solución

Ejercicio 28. Considere las funciones
$$F(x) = \int_0^{g(x)} f(x+t, xt)dt, \quad G(x,y) = \int_0^{y^2} g(x,y,z)dz$$
Calcule $F'(x)$ y $\frac{\partial G}{\partial y}(x,y)$

Ejercicio 29. Considere la función

$$
\begin{aligned}
u(x,t) &= \frac{1}{2}[f(x+ct)+f(x-ct)] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)ds \\
&+ \frac{1}{2c}\int_0^t \int_{x-c(t-s)}^{x+c(t-s)} F(\sigma,s)d\sigma ds
\end{aligned}
$$

Muestre que

$$
\begin{aligned}
\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} &= F(x,t) \\
u(x,0) &= f(x) \\
\frac{\partial u}{\partial t}(x,0) &= g(x)
\end{aligned}
$$

Ejercicio 30. Se desea resolver la ecuación de ondas

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c}\frac{\partial^2 \phi}{\partial t^2} = 0$$

donde c es una constante distinta de 0.
Haga el cambio de variables $\xi = x+ct$ y $\eta = x-ct$ de modo que se obtenga

$$\phi(x,t) = \psi(\xi(x,t),\eta(x,t))$$

donde $\psi(\xi,\eta)$ es la incógnita. En base a este cambio de variables demuestre que:

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c}\frac{\partial^2 \phi}{\partial t^2} = 4\frac{\partial^2 \psi}{\partial\eta\partial\xi}(\xi(x,y),\eta(x,y))$$

y en base a este resultado demuestre que toda solución ϕ de clase $\mathcal{C}^2$ de la ecuación de ondas se escribe

$$\phi(x,t) = f(\xi) + g(\eta)$$

con f y g funciones de variable real.

Cambio de variables

Ejercicio 31. Sea B_n la bola unitaria en $\mathbb{R}^n$. Muestre que

$$Vol(B_4) = 2\int_{B_3}\sqrt{1-(x^2+y^2+z^2)}dxdydz.$$

Tomando coordenadas esféricas muestre que el volumen de B_4 es igual a

$$2\int_{-\pi/2}^{\pi/2}\int_0^{2\pi}\int_0^1 r^2\operatorname{sen}(\phi)\sqrt{1-r^2}drd\theta d\phi$$

y concluya que $Vol(B_4) = \pi^2/2$. En general muestre que el volumen de la bola de radio r es $r^4\pi^2/2$.

Ejercicio 32. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ un difeomorfismo de clase $\mathcal{C}^1$ (es decir, un homeomorfismo con f y f^{-1} de clase $\mathcal{C}^1$), que satisface que $f(B) \subset B$, donde B es la bola unitaria cerrada y $|\det f'(x)| < 1$ para todo $x \in B$. Demuestre que para toda función continua $g : B \to \mathbb{R}^n$ se tiene que

$$\lim_{k\to\infty} \int_{f^k(B)} g(x)dx = 0$$

donde f^k es f compuesta consigo misma k veces.

Ejercicio 33. Calcular $\int_S \int dxdy$ donde S es el dominio limitado por

$$\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^2 = \frac{x^2}{h^2} - \frac{y^2}{k^2}$$

Ejercicio 34. Calcule

$$\iint_S x^2ydxdz + x^2zdxdy + x^3dydz$$

Donde S es la región que describe el volumen del cilindro $x^2 + y^2 = a^2$ entre y los planos $z = 0$ y $z = 2$.

Ejercicio 35. Describa el volumen del sólido M acotado por los paraboloides $z = x^2 + y^2$ y $z = 5(x^2 + y^2)$ y los planos de ecuaciones $z = 1$ y $z = 5$. Calcule el volumen.

Ejercicio 36. Calcular

$$\iint_S \sqrt[3]{1 - \left(\frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{25}\right)}dxdydz$$

donde $S = \frac{x^2}{9} + \frac{y^2}{16} + \frac{z^2}{25} \leq 4$.

Ejercicio 37. Calcule el volumen de la región encerrada dentro de la esfera $x^2 + y^2 + z^2 = 5$ y el cilindro $(x - 1)^2 + y^2 = 1$.

Indicación. Use $x = \arccos(\varphi)$ e $y = \operatorname{sen}(\varphi)$.

CAPÍTULO 4

Topología Básica

La idea de este capítulo es formalizar conceptos abstractos que son de alta importancia en los temas que siguen. No es posible entender a cabalidad los temas que vienen a continuación en este apunte sin tener claro lo más básico de topología, por esto es que preferimos destinar espacio a estos conceptos.

4.1. Normas y espacios normados

Definición 4.1. Sea E un espacio vectorial. Una norma en E es una función que satisface las siguientes propiedades:

$$\|\cdot\| : E \to \mathbb{R}_+ \cup \{0\}$$

1. Positividad: $\|\boldsymbol{x}\| = 0 \Leftrightarrow \boldsymbol{x} = 0$
2. Linealidad: $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\| \ \forall \lambda \in \mathbb{R}, \forall \boldsymbol{x} \in E$
3. Desigualdad triangular: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \ \forall \boldsymbol{x}, \boldsymbol{y} \in E$

Al par ordenado $(E, \|\cdot\|)$ se lo conoce como espacio vectorial normado.

Ejemplo 4.1. En $\mathbb{R}^n$ podemos definir muchas normas:

1. Norma 1: $\|\boldsymbol{x}\|_1 = |x_1| + \ldots + |x_n|$
2. Norma 2 o euclideana: $\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + \ldots + x_n^2}$
3. Norma infinito o uniforme: $\|\boldsymbol{x}\|_\infty = \max\limits_{i=1,2,\ldots,n} |x_i|$
4. Norma ρ: $\|\boldsymbol{x}\|_\rho = \sqrt[\rho]{|x_1|^\rho + \ldots + |x_n|^\rho}$ para $1 \leq \rho < \infty$

De forma que $(\mathbb{R}^n, \|\cdot\|)$ con cualquiera de las normas definidas tiene estructura de espacio vectorial normado.

Demostrar que $\|\cdot\|_2, \|\cdot\|_1$ y $\|\cdot\|_\infty$ satisfacen efectivamente las propiedades de una norma es tarea sencilla. Un poco más difícil es demostrar que $\|\cdot\|_p$ con $1 < p < \infty$ es también una norma.

Las normas mencionadas son casos particulares de $\|\boldsymbol{x}\|_\rho$. Las normas 1, 2 e infinito se obtienen cuando ρ toma los valores 1, 2 y tendiendo a infinito respectivamente.

Ejemplo 4.2. Para fijar ideas, en $\mathbb{R}^2$ el dibujo de $\|\boldsymbol{x}\|_\rho = 1$ con ρ igual a 1/2, 1, 2 y tendiendo a infinito nos queda de la siguiente forma

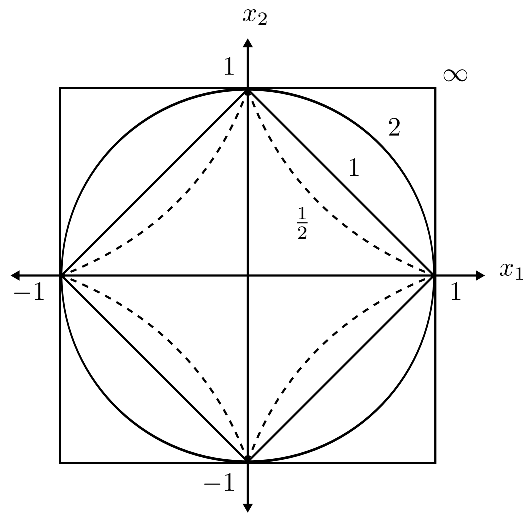


Figura 4.1: Normas en $\mathbb{R}^2$

Decimos $\|\boldsymbol{x}\|_\rho$ es norma cuando $\rho \geq 1$ pues en caso de que $\rho < 1$ no se cumple la desigualdad triangular.

Sean $\boldsymbol{x} = (1,0)$ e $\boldsymbol{y} = (0,1)$. Si $p = 1/2$ tenemos que

$$\begin{aligned}
\|\boldsymbol{x}\|_\rho &= (|0|^{1/2} + |1|^{1/2})^2 = 1 \\
\|\boldsymbol{y}\|_\rho &= (|1|^{1/2} + |0|^{1/2})^2 = 1
\end{aligned}$$

de lo cual se obtiene $\|\boldsymbol{x}\|_\rho + \|\boldsymbol{y}\|_\rho = 2$ mientras que $\|\boldsymbol{x}+\boldsymbol{y}\|_\rho = 4$.

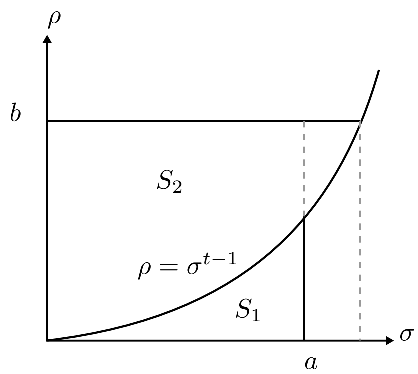
Lema 4.1. *Sean $a, b \geq 0 \in \mathbb{R}$ y $\rho, \sigma > 1 \in \mathbb{R}$ tales que $\frac{1}{\rho} + \frac{1}{\sigma} = 1$. Entonces,*

$$\frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma} \geq ab \tag{4.1}$$

Demostración. Consideremos la siguiente curva

$$\rho(\sigma) = \sigma^{t-1} \Leftrightarrow \sigma(\rho) = \rho^{q-1}$$

la gráfica de esto corresponde a lo siguiente:



Consideremos las áreas S_1 y S_2 del gráfico y se tiene que

$$S_1 = \int_0^a \sigma^{t-1} d\sigma = \frac{a^\rho}{\rho} \ , \ S_2 = \int_0^b \rho^{q-1} d\rho = \frac{b^\sigma}{\sigma}$$

De la figura se puede inferir que

$$\begin{aligned} ab &\leq S_1 + S_2 \\ ab &\leq \frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma} \end{aligned}$$

Una forma más rigurosa de demostrar este lema es por medio de la definición de función convexa. Una función $f : A \subseteq \mathbb{R} \to \mathbb{R}$ es convexa si

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \qquad \forall x, y \in A, 0 \leq \lambda \leq 1$$

Otro criterio para determinar la convexidad de una función de una variable, en caso de que sea dos veces continua y diferenciable, es determinar que su segunda derivada sea no decreciente ($f''(x) \geq 0 \ \forall x \in A$).

Tomando la función $-\ln(x)$ (verifique la convexidad mediante la segunda derivada), escojamos $\lambda = 1/\rho$. Entonces por convexidad

$$\begin{aligned} -\ln\left(\frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma}\right) &\leq -\frac{\ln(a^\rho)}{\rho} - \frac{\ln(b^\sigma)}{\sigma} \\ -\ln\left(\frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma}\right) &\leq -\ln(ab) \\ \ln\left(\frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma}\right) &\geq \ln(ab) \\ \frac{a^\rho}{\rho} + \frac{b^\sigma}{\sigma} &\geq ab \qquad (4.2) \end{aligned}$$

■

Teorema 4.1. $\|\cdot\|_\rho$ *es una norma.*

Demostración. Debemos demostrar las tres propiedades que aparecen en la definición 4.1.

Positividad: $\|\boldsymbol{x}\|_\rho \geq 0 \ \forall \boldsymbol{x} \in \mathbb{R}^n$
Primero veamos qué pasa con $\|\boldsymbol{x}\|_\rho \geq 0$

$$|x_i| \geq 0 \ \forall x_i \in \mathbb{R} \Rightarrow \|\boldsymbol{x}\|_\rho = \left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} \geq 0$$

Luego debe cumplirse que $\|\boldsymbol{x}\|_\rho = 0 \Leftrightarrow \boldsymbol{x} = 0$, entonces si $\boldsymbol{x} = 0$

$$|x_i| = 0 \ \forall x_i \in \mathbb{R} \Rightarrow \|\boldsymbol{x}\|_\rho = \left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} = 0$$

la otra implicancia es como sigue, si $\|\boldsymbol{x}\|_\rho = 0$

$$\left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} = (|x_1|^\rho + \ldots + |x_n|^\rho)^{1/\rho} = 0$$

luego $|x_i| \in \mathbb{R}_+$ y, por definición, $|x_i| \geq 0$ además de que la suma de elementos positivos es nula sólo si todos los elementos son nulos

$$\left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} = |x_1|^\rho + \ldots + |x_n|^\rho = 0 \Rightarrow \boldsymbol{x} = 0$$

Linealidad: $\|\lambda\boldsymbol{x}\|_\rho = |\lambda|\|\boldsymbol{x}\|_\rho \ \forall \boldsymbol{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}$

$$\begin{aligned}
\|\lambda\boldsymbol{x}\|_\rho &= \left(\sum_{i=1}^n |\lambda x_i|^\rho\right)^{1/\rho} \\
&= \left(|\lambda|^\rho \sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} \\
&= |\lambda|\|\boldsymbol{x}\|_\rho
\end{aligned}$$

Desigualdad triangular: $\|\boldsymbol{x}+\boldsymbol{y}\|_\rho \leq \|\boldsymbol{x}\|_\rho + \|\boldsymbol{y}\|_\rho \ \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$\left(\sum_{i=1}^n |x_i + y_i|^\rho\right)^{1/\rho} \leq \left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} + \left(\sum_{i=1}^n |y_i|^\rho\right)^{1/\rho}$$

Sea $\rho = 1$

$$\sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|)$$

Sean $\rho > 1$ y $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Si tomamos un real $x \geq 0$ entonces $x \in \mathbb{R}_+$ y se tiene que $x = |x|$. Usando esta propiedad, reescribiremos (4.1) Escojamos $a = \frac{|x_i|}{(\sum_{j=1}^n |x_j|^\rho)^{1/\rho}}$ y $b = \frac{|y_i|}{(\sum_{j=1}^n |y_j|^\sigma)^{1/\sigma}}$, entonces la desigualdad nos queda

$$\frac{|x_i|}{(\sum_{j=1}^n |x_j|^\rho)^{1/\rho}} \cdot \frac{|y_i|}{(\sum_{j=1}^n |y_j|^\sigma)^{1/\sigma}} \leq \frac{1}{\rho} \cdot \left(\frac{|x_i|}{(\sum_{j=1}^n |x_j|^\rho)^{1/\rho}}\right)^\rho + \frac{1}{\sigma} \cdot \left(\frac{|y_i|}{(\sum_{j=1}^n |y_j|^\sigma)^{1/\sigma}}\right)^\sigma$$

$$\frac{|x_i y_i|}{(\sum_{j=1}^n |x_j|^\rho)^{1/\rho} \cdot (\sum_{j=1}^n |y_j|^\sigma)^{1/\sigma}} \leq \frac{1}{\rho} \cdot \frac{|x_i|^\rho}{\sum_{j=1}^n |x_j|^\rho} + \frac{1}{\sigma} \cdot \frac{|y_i|^\sigma}{\sum_{j=1}^n |y_j|^\sigma}$$

Aplicando $\sum_{i=1}^n(\cdot)$ a la última desigualdad obtenemos

$$\begin{aligned}
\frac{\sum_{i=1}^n |x_i y_i|}{(\sum_{j=1}^n |x_j|^\rho)^{1/\rho} \cdot (\sum_{j=1}^n |y_j|^\sigma)^{1/\sigma}} &\leq \left(\frac{1}{\rho} + \frac{1}{\sigma} = 1\right) \\
\sum_{i=1}^n |x_i y_i| &\leq \left(\sum_{j=1}^n |x_j|^\rho\right)^{1/\rho} \cdot \left(\sum_{j=1}^n |y_j|^\sigma\right)^{1/\sigma} \qquad (4.3)
\end{aligned}$$

Como $|x_i + y_i| \leq |x_i| + |y_i|$

$$\begin{aligned}
(|x_i| + |y_i|)^\rho &= (|x_i| + |y_i|)^{\rho-1}(|x_i| + |y_i|) \\
|x_i + y_i|^\rho &\leq (|x_i| + |y_i|)^{\rho-1}(|x_i| + |y_i|)
\end{aligned}$$

Aplicando $\sum_{i=1}^n(\cdot)$ a la última desigualdad obtenemos

$$\begin{aligned}
\sum_{i=1}^n |x_i + y_i|^\rho &\leq \sum_{i=1}^n (|x_i| + |y_i|)^{\rho-1}(|x_i| + |y_i|) \\
\sum_{i=1}^n |x_i + y_i|^\rho &\leq \sum_{i=1}^n (|x_i| + |y_i|)^{\rho-1}|x_i| + \sum_{i=1}^n (|x_i| + |y_i|)^{\rho-1}|y_i|
\end{aligned}$$

Combinando esto último con la ecuación (4.3) y como $\sigma(\rho - 1) = \rho$ se tiene

$$\begin{aligned}
\sum_{i=1}^n |x_i + y_i|^\rho &\leq \left[\left(\sum_{i=1}^n |x_i|^\rho\right)^{\frac{1}{\rho}} + \left(\sum_{i=1}^n |y_i|^\rho\right)^{\frac{1}{\rho}}\right] \left(\sum_{i=1}^n (|x_i| + |y_i|)^{\sigma(\rho-1)}\right)^{\frac{1}{\sigma}} \\
\sum_{i=1}^n |x_i + y_i|^\rho &\leq \left[\left(\sum_{i=1}^n |x_i|^\rho\right)^{\frac{1}{\rho}} + \left(\sum_{i=1}^n |y_i|^\rho\right)^{\frac{1}{\rho}}\right] \left(\sum_{i=1}^n (|x_i| + |y_i|)^{\rho}\right)^{\frac{1}{\sigma}} \\
\left(\sum_{i=1}^n |x_i + y_i|^\rho\right)^{1-\frac{1}{\sigma}} &\leq \left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} + \left(\sum_{i=1}^n |y_i|^\rho\right)^{1/\rho} \\
\left(\sum_{i=1}^n |x_i + y_i|^\rho\right)^{1/\rho} &\leq \left(\sum_{i=1}^n |x_i|^\rho\right)^{1/\rho} + \left(\sum_{i=1}^n |y_i|^\rho\right)^{1/\rho}
\end{aligned}$$

$\blacksquare$

Ejemplo 4.3. Si $E = \{f : [a,b] \to \mathbb{R} : f \text{ es continua } \} = C[a,b]$ entonces definimos

$$\|\cdot\|_\infty : E \to \mathbb{R} \cup \{0\}$$

$$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$$

Ejemplo 4.4. En $\mathcal{P} = \{\text{polinomios de } \mathbb{R}\}$ definimos

$$\|p\| = |p(1)| + |p(0)| + |p(-1)|$$

lo anterior no es una norma pues no satisface la condición 1 de norma. En los polinomios de grado dos si es norma.

En un espacio vectorial E podemos considerar diferentes normas.

Teorema 4.2. (Equivalencia de normas)
Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ en E se dicen equivalentes si existen constantes c_1 y c_2 no negativas tales que:

$$c_1\|\boldsymbol{x}\|_1 \leq \|\boldsymbol{x}\|_2 \leq c_2\|\boldsymbol{x}\|_1 \ \forall \boldsymbol{x} \in E$$

Demostración. Supongamos que $\|\cdot\|_1$ y $\|\cdot\|_2$ en E son equivalentes.

Tomemos el lado izquierdo de la desigualdad y supongamos que no existe $c_1 \in \mathbb{R}_+$ que cumpla $c_1\|\boldsymbol{x}\|_1 \leq \|\boldsymbol{x}\|_2$. Entonces

$$\inf_{\boldsymbol{x}\in E\setminus\{0\}} \frac{\|\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_1} = 0$$

Escojamos $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ en E tal que $\{\boldsymbol{x}_n\} \to 0$, entonces

$$\frac{\|\boldsymbol{x}_n\|_2}{\|\boldsymbol{x}_n\|_1} \to 0$$

Consideremos que $\|\boldsymbol{x}_n\|_1, \|\boldsymbol{x}_n\|_2 \in \mathbb{R}_+$ por lo que esto último lo podemos reescribir como

$$\left\|\frac{\boldsymbol{x}_n}{\|\boldsymbol{x}\|_1}\right\|_2 \to 0$$

Definamos $\boldsymbol{y}_n = \frac{\boldsymbol{x}_n}{\|\boldsymbol{x}\|_1}$, entonces $\|\boldsymbol{y}_n\|_1 = 1$ e $\|\boldsymbol{y}_n\|_2 \to 0$ lo que significa que las normas no pueden ser equivalentes.

Tomemos el lado derecho de la desigualdad y supongamos que no existe $c_2 \in \mathbb{R}_+$ que cumpla $\|\boldsymbol{x}\|_2 \leq c_2\|\boldsymbol{x}\|_1$. Entonces

$$\inf_{\boldsymbol{x}\in E\setminus\{0\}} \frac{\|\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_2} = 0$$

Escojamos $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ en E tal que $\{\boldsymbol{x}_n\} \to 0$, entonces

$$\frac{\|\boldsymbol{x}_n\|_1}{\|\boldsymbol{x}_n\|_2} \to 0$$

Consideremos que $\|\boldsymbol{x}_n\|_1, \|\boldsymbol{x}_n\|_2 \in \mathbb{R}_+$ por lo que esto último lo podemos reescribir como

$$\left\|\frac{\boldsymbol{x}_n}{\|\boldsymbol{x}\|_2}\right\|_1 \to 0$$

Definamos $\boldsymbol{y}_n = \frac{\boldsymbol{x}_n}{\|\boldsymbol{x}\|_2}$, entonces $\|\boldsymbol{y}_n\|_1 \to 0$ y $\|\boldsymbol{y}_n\|_2 = 1$ lo que significa que las normas no pueden ser equivalentes.

■

Cuando a la estructura de espacio vectorial se le agrega una norma, se le dota de propiedades topológicas. Veremos más adelante que si $\|\cdot\|_1$ y $\|\cdot\|_2$ son normas equivalentes en E entonces $(E, \|\cdot\|_1)$ y $(E, \|\cdot\|_2)$ tienen las mismas propiedades topológicas.

Definición 4.2. Dados dos puntos $\boldsymbol{x}, \boldsymbol{y} \in E$ definimos la distancia entre $\boldsymbol{x}$ e $\boldsymbol{y}$ por:

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$$

Observamos que la noción de distancia entre dos puntos depende de la norma considerada.

Ejemplo 4.5. En $M_{n\times m}(\mathbb{R})$ definimos dos normas

$$\|A\|_\infty = \max_{i,j} |a_{ij}|$$

$$\|A\|_1 = \sum_i \sum_j |a_{ij}|$$

Si $A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$ y $B = \begin{pmatrix} 6 & 7 \\ 8 & 9 \end{pmatrix}$

$$\|A-B\|_\infty = \left\| \begin{pmatrix} -4 & -4 \\ -4 & -4 \end{pmatrix} \right\|_\infty = 4 \qquad , \qquad \|A-B\|_1 = \left\| \begin{pmatrix} -4 & -4 \\ -4 & -4 \end{pmatrix} \right\|_1 = 16$$

4.2. Sucesiones

En el estudio de la topología de un espacio normado, un rol muy importante es jugado por las sucesiones. Como en el caso de $\mathbb{R}$ uno puede definir sucesiones de vectores en $\mathbb{R}^n$.

Definición 4.3. (Sucesión)
Una sucesión en el espacio vectorial E es una función

$$\begin{array}{ccc} \mathbb{N} & \to & E \\ n & \mapsto & \boldsymbol{x}_n \end{array}$$

y se anota usualmente como $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$. Una sucesión $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$ si

$$\forall \varepsilon > 0,\ \exists n \in \mathbb{N} \text{ tal que } \|\boldsymbol{x}_n - \boldsymbol{x}\| < \varepsilon,\ \forall n \geq N$$

Si una sucesión converge a un vector $\boldsymbol{x}$ diremos que este es el límite de la sucesión y se denota

$$\lim_n \boldsymbol{x}_n = \boldsymbol{x}$$

Proposición 4.1. El límite de una sucesión, cuando existe, es único.

Demostración. Supongamos que $\lim \boldsymbol{x}_n = \boldsymbol{x}_1$ y $\lim \boldsymbol{x}_n = \boldsymbol{x}_2$. Tendremos que para todo $\varepsilon > 0$ existen enteros n_1 y n_2 tales que $\|\boldsymbol{x}_n - \boldsymbol{x}\| \leq \varepsilon/2$ para todo $n \geq n_1$ y $\|\boldsymbol{x}_n - \boldsymbol{x}\| \leq \varepsilon/2$ para todo $n \geq n_2$. Sea $m = \max\{n_1, n_2\}$ tenemos que

$$\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leq \|\boldsymbol{x}_1 - \boldsymbol{x}_m\| + \|\boldsymbol{x}_m - \boldsymbol{x}_2\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

y esta desigualdad es válida para todo $\varepsilon > 0$, por lo tanto $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = 0$ y por las propiedades de la norma concluimos que $\boldsymbol{x}_1 = \boldsymbol{x}_2$. ∎

Nota 4.1. En adelante escribiremos $\boldsymbol{x}_n \to \boldsymbol{x}$ en lugar de $\lim_n \boldsymbol{x}_n = \boldsymbol{x}$ para referirnos al límite de una sucesión.

4.2.1. Sucesiones de Cauchy

Definición 4.4. (Sucesión de Cauchy)
Una sucesión $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ se dice de Cauchy si $\forall\varepsilon > 0$ existe $n \in \mathbb{N}$ tal que

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \varepsilon \ \forall n, m \geq N$$

Nota 4.2. Las nociones de sucesiones convergentes y sucesiones de Cauchy no cambian según la norma por la equivalencia de estas.

Proposición 4.2. Toda sucesión convergente es sucesión de Cauchy.

Demostración. Supongamos que $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$. Dado $\varepsilon > 0$ existirá $n_0 \in \mathbb{N}$ tal que $\|\boldsymbol{x}_n - \boldsymbol{x}\| \leq \varepsilon/2$ para cualquier $n \geq n_0$. Entonces

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| \leq \|\boldsymbol{x}_n - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{x}_m\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \ \forall n, m \geq n_0$$

lo cual concluye la demostración. ■

Nota 4.3. En general no toda sucesión de Cauchy es convergente por lo que la recíproca de la proposición es falsa.

Ejemplo 4.6. En $\mathbb{Q}^2$ usamos la norma euclideana. La sucesión definida por

$$\boldsymbol{x}_k = \left(\left(1 + \frac{1}{k}\right)^k, \frac{1}{k}\right)$$

es de Cauchy pero no converge en $\mathbb{Q}^2$.

Ejemplo 4.7. En $C[-1,1]$ dotado de la norma

$$\|f\|_1 = \int_{-1}^{1} |f(x)| dx$$

consideramos la sucesión

$$f_n(x) = \begin{cases} 1 - (1-x)^n & \text{si } x > 0 \\ -1 + (1+x)^n & \text{si } x \leq 0 \end{cases}$$

Supongamos sin perdida de generalidad que $n \geq m$

$$\begin{aligned} \|f_n - f_m\|_1 &= \int_0^1 |(1-x)^n - (1-x)^m| dx + \int_{-1}^0 |(1+x)^n - (1+x)^m| dx \\ &= \int_0^1 (1-x)^m |(1-x)^{n-m} - 1| dx + \int_{-1}^0 (1+x)^m |(1+x)^{n-m} - 1| \\ &\leq \int_0^1 (1+x)^m dx + \int_{-1}^0 (1+x)^m dx \leq \frac{2}{m+1} \end{aligned}$$

es decir,

$$\|f_n - f_m\|_1 \leq \frac{2}{\min\{n, m\} + 1}$$

por lo tanto $\{f_n\}$ es de Cauchy. Pero $\{f_n\}$ no tiene limite en $C([-1,1])$ En efecto, supongamos que existe la función límite que llamaremos f. Entonces

$$\|f_n - f\|_1 \overset{n\to\infty}{\to} 0$$

Como

$$\int_a^1 |f_n - f|dx \leq \int_{-1}^1 |f_n - f|dx$$

entonces

$$\int_a^1 |f_n - f|dx \overset{n\to\infty}{\to} 0 \ \forall a \in [-1,1]$$

tomemos $a \in (0,1]$. Como en $[a,1]$, f_n converge uniformemente a 1, entonces

$$\int_a^1 |1 - f|dx = 0 \Rightarrow f = 1 \text{ en } [a,1] \ \forall a \in (0,1]$$

por lo tanto $f(x) = 1 \ \forall x \in (0,1]$. Análogamente $f(x) = -1 \ \forall a \in [-1,0)$
Concluimos así que f no puede ser continua.

4.3. Espacios de Banach

Definición 4.5. (Espacio de Banach)
Un espacio vectorial normado E se dice espacio de Banach si toda sucesión de Cauchy en E converge en E, es decir

$$\{\boldsymbol{x}_n\}_{n\in\mathbb{N}} \subseteq E, \text{ es de Cauchy } \Rightarrow \exists\, \boldsymbol{x} \in E \text{ tal que } \boldsymbol{x}_n \overset{n\to\infty}{\to} \boldsymbol{x}$$

en este caso E también recibe el nombre de espacio vectorial completo.

Nota 4.4. La gran mayoría de los modelos que se utilizan en modelos matemáticos de la ingeniería se construyen sobre espacios de Banach.

Ejemplo 4.8. $(\mathbb{R}, |\cdot|)$. En $\mathbb{R}$ toda sucesión de Cauchy es convergente. Esto es una consecuencia del axioma del supremo.

Ejemplo 4.9. $(\mathbb{R}^n, \|\cdot\|_2)$ es un espacio de Banach. En efecto, si $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$ es una sucesión de Cauchy, entonces

$$\forall\, \varepsilon > 0 \ \exists N \in \mathbb{N} \text{ tal que } \|\boldsymbol{x}_k - \boldsymbol{x}_m\|_2 < \varepsilon \ \forall k, l > N$$

lo que implica que

$$|x_k^i - x_m^i| < \varepsilon \ \forall k, m > N \ \forall i \in \{1, \ldots, n\}$$

donde x_k^i es la i-ésima componente de $\boldsymbol{x}_k$. Por lo tanto la sucesión $\{x_k^i\}_{k\in\mathbb{N}}$ es de Cauchy en $\mathbb{R}$ y en consecuencia converge. Denotemos por x^i su límite. Tenemos asi que dado $\varepsilon > 0$ existe N_i que satisface

$$|x_k^i - x^i| < \frac{\varepsilon}{\sqrt{n}} \ \forall k > N_i$$

escogiendo $N = \text{máx}\{N_i : i = 1, \ldots, n\}$ y llamando $\boldsymbol{x}$ al vector $(x^1, \ldots, x^n)$ tenemos finalmente que

$$\|\boldsymbol{x}_k - \boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^n |x_k^i - x^i|^2} < \varepsilon \ \forall k > N$$

es decir la sucesión $\boldsymbol{x}_k$ converge a $\boldsymbol{x}$ en $\mathbb{R}^n$

Nota 4.5. Es fácil ver que

$$\boldsymbol{x}_k \to \boldsymbol{x} \text{ en } (\mathbb{R}^n, \|\cdot\|_2) \Leftrightarrow x_k^i \to x^i \ \forall i \in \{1,..,n\}$$

Ejemplo 4.10. $(M_{n\times m}(\mathbb{R}), \|\cdot\|_\infty)$ es completo. Sea $\{A_k\}_{k\in\mathbb{N}}$ una sucesión de Cauchy. Entonces $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ tal que

$$\|A_k - A_m\|_\infty < \varepsilon \ \forall k, m > N$$

es decir,

$$\max_{\substack{1\le i\le n\\ 1\le j\le m}} |a_{ij}^k - a_{ij}^m| < \varepsilon$$

en consecuencia cada una de las sucesiones $\{a_{ij}^k\}_{k\in\mathbb{N}}$ son de Cauchy en $\mathbb{R}$. Cada una de ellas converge entonces a un límite que denotamos a_{ij}. De esta manera, dado $\varepsilon > 0\ \exists N_{ij} > 0$ tal que

$$|a_{ij}^k - a_{ij}| < \varepsilon \ \forall k > N_{ij}$$

escogiendo $N = \max\{N_{ij} \text{ tal que } i = 1,\dots,n,\ j = 1,\dots,m\}$, obtenemos

$$\max_{\substack{1\le i\le n\\ 1\le j\le m}} |a_{ij}^k - a_{ij}| < \varepsilon \ \forall k > N$$

o sea, si $A = (a_{ij})$

$$\|A_k - A\|_\infty < \varepsilon \ \forall k > n$$

por lo tanto $A_k \to A$. El espacio de matrices de dimensión $n\times m$ con la norma $\|\cdot\|_\infty$ es un espacio de Banach.

Ejemplo 4.11. $(C([a,b],\mathbb{R}), \|\cdot\|_\infty)$ es un espacio de Banach. Sea $\{f_n\}_{n\in\mathbb{N}}$ una sucesión de Cauchy. Dado $\varepsilon > 0$, $\exists N \in \mathbb{N}$ tal que

$$\|f_k - f_m\|_\infty = \sup_{x\in[a,b]} |f_k(x) - f_m(x)| < \varepsilon \ \forall k, m > N$$

en particular, para cualquier $x \in [a,b]$ se tendrá que

$$|f_k(x) - f_m(x)| < \varepsilon \ \forall k, m > N$$

es decir, $\forall x \in [a,b]$ la sucesión $\{f_n(x)\}_{n\in\mathbb{N}}$ es de Cauchy en $\mathbb{R}$, y por lo tanto converge a un límite que llamaremos $f(x)$

$$\lim_{k\to\infty} f_k(x) = f(x)$$

De esta manera, hemos definido una función $f : [a,b] \to \mathbb{R}$. Probaremos que esta función es el límite de la sucesión $\{f_n\}$ en $C([a,b],\mathbb{R})$.

Teníamos que dado $\varepsilon > 0$, $\exists N > 0$ tal que

$$|f_k(x) - f_m(x)| < \varepsilon \ \forall k, m > N \ \forall x \in [a,b]$$

lo que implica que

$$|f_k(x) - f(x)| < \varepsilon \ \forall k > N \ \forall x \in [a,b]$$

y entonces

$$\|f_k - f\|_\infty < \varepsilon \ \forall k > N.$$

Con esto hemos probado que

$$f_n \overset{\|\cdot\|_\infty}{\to} f$$

es decir, f es el límite uniforme de las funciones continuas f_n. Para concluir la demostración de que $(C([a,b],\mathbb{R}), \|\cdot\|_\infty)$ es Banach debemos probar que f es continua.

Proposición 4.3. El límite uniforme de funciones continuas de la forma $f : [a, b] \to \mathbb{R}$ es una función continua.

Demostración. Sea $\{f_n\}_{n\in\mathbb{N}}$ una sucesión de funciones continuas en $[a, b]$ que converge uniformemente a una función f. Sea $x_0 \in [a, b]$ y $\varepsilon > 0$, entonces existe N tal que

$$\|f_k - f\|_\infty < \frac{\varepsilon}{3} \ \forall k \geq N$$

en particular tendremos que

$$|f_N(x) - f(x)| < \frac{\varepsilon}{3} \ \forall x \in [a, b]$$

pero f_N es una función continua y por lo tanto $\exists \delta > 0$ tal que si $|x - x_0| < \delta$ entonces

$$|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3}$$

Con todo esto obtenemos que si $|x - x_0| < \delta$

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

Es decir, f es una función continua. ■

Con esto concluimos que $(C([a, b], \mathbb{R}), \|\cdot\|_\infty)$ es un Espacio de Banach.

4.4. Subsucesiones

Definición 4.6. (Subsucesión)
Sea $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ una sucesión en un espacio normado $(E, \|\cdot\|)$. Consideremos una función $f : \mathbb{N} \to \mathbb{N}$ estrictamente creciente, es decir, $n < m \Rightarrow f(n) < f(m)$. Entonces, la nueva sucesión $\{\boldsymbol{x}_{f(k)}\}_{k\in\mathbb{N}}$ se llama subsucesión de $\{\boldsymbol{x}_n\}$. A menudo se anota $n_k = f(k)$ y así la subsucesión se anota como $\{\boldsymbol{x}_{n_k}\}_{k\in\mathbb{N}}$.

Ejemplo 4.12. Las sucesiones siguientes

$$\left\{\left(-\frac{1}{2}\right)^{2n}\right\}_{n\in\mathbb{N}}, \ \left\{\left(-\frac{1}{2}\right)^{2n+1}\right\}_{n\in\mathbb{N}} \text{ y } \left\{\left(-\frac{1}{2}\right)^{8n+7}\right\}_{n\in\mathbb{N}},$$

son subsucesiones de $\{(-\frac{1}{2})^n\}_{n\in\mathbb{N}}$

Ejemplo 4.13. La sucesión $\{\boldsymbol{x}_n\} \subseteq \mathbb{R}^2$ definida por

$$\boldsymbol{x}_n = \left((-1)^n, \left(1 + \frac{1}{n}\right)^n\right)$$

no converge. Sin embargo la subsucesión $\{\boldsymbol{x}_{2n}\}_{n\in\mathbb{N}}$ si converge y lo hace a $(1, e) \in \mathbb{R}^2$. La subsucesión $\{\boldsymbol{x}_{2n+1}\}_{n\in\mathbb{N}}$ también converge, pero a $(-1, e) \in \mathbb{R}^2$. Por otra parte la subsucesión $\{\boldsymbol{x}_{3n}\}_{n\in\mathbb{N}}$ no converge (nótese como cambia el signo de $(-1)^{kn}$) para los valores $k = 1, 2, 3$).

Ejemplo 4.14. La sucesión $\boldsymbol{x}_n = (2^n, \frac{1}{n}, \frac{n}{n+1}) \in \mathbb{R}^3$ no converge. En este caso $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ no tiene subsucesiones convergentes.

Ejemplo 4.15. En el espacio de Banach $(C([0,1],\mathbb{R}),\|\cdot\|_\infty)$ consideremos la siguiente sucesión:

$$f_n(x) = \begin{cases} 0 & ,0 \leq x \leq \frac{1}{n+1} \\ 1+(n+1)(nx-1) & ,\frac{1}{n+1} \leq x \leq \frac{1}{n} \\ 1-(n-1)(nx-1) & ,\frac{1}{n} \leq x \leq \frac{1}{n-1} \\ 0 & ,\frac{1}{n-1} \leq x \leq 1 \end{cases}$$

Para esta sucesión se cumple que:

- $\|f_n\|_\infty = 1 \ \forall n \in \mathbb{N}$
- $\|f_n - f_{n+1}\|_\infty = 1 \ \forall n \in \mathbb{N}$
- $\|f_n - f_k\|_\infty = 1 \ \forall n \neq k$

Entonces $\{f_n\}$ no converge y ninguna subsucesión puede hacerlo, pues no son de Cauchy.
Esto muestra que en $(C([0,1],\mathbb{R}),\|\cdot\|_\infty)$ hay sucesiones acotadas que no tienen subsucesiones convergentes. En particular mostramos que $B = \overline{B}(0,1)$, la bola unitaria cerrada, no es compacta.

Teorema 4.3. *Sea $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ una sucesión en un espacio normado. Entonces $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$ si y sólo si toda subsucesión $\{\boldsymbol{x}_{n_k}\}_{n\in\mathbb{N}}$ de $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$*

Demostración.
($\Leftarrow$): Directo pues $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ es una subsucesión $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$.

($\Rightarrow$): $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$, entonces $\forall\varepsilon > 0\ \exists N \in \mathbb{N}$ tal que

$$\|\boldsymbol{x}_n - \boldsymbol{x}\| < \varepsilon \ \forall n \geq N$$

Sea $\{\boldsymbol{x}_{n_k}\}_{n\in\mathbb{N}}$ una subsucesión de $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ entonces $\exists k \in \mathbb{N}$ tal que $n_1 < n_2 < \ldots < n_k < n_{k+1} < n_{k+2} < \ldots$. Entonces

$$\|\boldsymbol{x}_{n_k} - \boldsymbol{x}\| < \varepsilon \ \forall k \geq K$$

Por lo tanto $\{\boldsymbol{x}_{n_k}\}_{n\in\mathbb{N}}$ converge a $\boldsymbol{x}$. ■

El siguiente es un teorema fundamental en la topología de $\mathbb{R}$ que puede extenderse a $\mathbb{R}^n$.

Teorema 4.4. (Teorema de Bolzano-Weierstrass en $\mathbb{R}$)
Sean $[a,b]$ un intervalo cerrado y acotado y $f:[a,b]\to\mathbb{R}$ continua tal que $f(a)$ y $f(b)$ tienen signos contrarios, entonces existe $c \in [a,b]$ tal que $f(c) = 0$.

Demostración. Sin pérdida de generalidad supongamos que $f(a) < 0$ y $f(b) > 0$. Escojamos $c \in (a,b)$ y de esto se tienen tres casos:

1. $f(c) < 0$ y nos restringimos al intervalo $[a_1,b_1]$ con $a_1 = c$ y $b_1 = b$.
2. $f(c) = 0$ en este caso concluye la demostración.
3. $f(c) > 0$ y nos restringimos al intervalo $[a_1,b_1]$ con $a_1 = a$ y $b_1 = c$.

Para los casos 1. y 3. consideremos intervalos $[a_n,b_n] \subset [a_{n-1},b_{n-1}] \subset \ldots \subset [a,b]$ tal que $f(a_n) < 0$ y $f(b_n) > 0$. Escojamos para cada intervalo un c que es punto medio y así cada intervalo es la mitad del anterior. De esta forma, $b_n - a_n = \frac{b-a}{2^n}$ y para $n\to\infty$ se tendrá que $a_n - b_n \to 0$ y $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

Sea $c = \lim_{n\to\infty} a_n$, por ser f continua

$$f(c) = f\left(\lim_{n\to\infty} a_n\right) = \lim_{n\to\infty} f(a_n)$$

como $f(a_n) < 0$ tenemos que $\lim_{n\to\infty} f(a_n) \leq 0$.
Análogamente si tomamos

$$f(c) = f\left(\lim_{n\to\infty} b_n\right) = \lim_{n\to\infty} f(b_n)$$

como $f(b_n) < 0$ tenemos que $\lim_{n\to\infty} f(a_n) \geq 0$. Se concluye entonces que $f(c) = 0$. ■

Teorema 4.5. (Teorema de Bolzano-Weierstrass en $\mathbb{R}^n$)
Toda sucesión acotada en $\mathbb{R}^n$ tiene a lo menos una subsucesión convergente.

Demostración. Sea $\{\boldsymbol{x}_n\}$ una sucesión acotada en $\mathbb{R}^n$. Para cada componente de la sucesión se tiene $\{x_n^i\}_{i\in\mathbb{N}}$ que corresponde a una sucesión acotada de números reales. De acuerdo al teorema 1.3, $\{x_n^i\}_{i\in\mathbb{N}}$ tiene a lo menos una subsucesión convergente $\{x_{\gamma(n)}^i\}_{i\in\mathbb{N}}$. Esto es, si $\{x_n^i\} \to a^i \in \mathbb{R}$ entonces $\{x_{\gamma(n)}^i\} \to b^i \in \mathbb{R}$.
Sea $A_1 \subset \mathbb{N}$ y $a_1^i \in A$, entonces $\lim_{n\in A_1}\{x_n^1\} = a_1^i$. En forma análoga, sea $A_2 \subset A_1$, entonces $\lim_{n\in A_2}\{x_n^2\} = a_2^i$. Podemos construir la inclusión $A_n \subset A_{n-1}$ tal que $A_n \subset \mathbb{N}$ y $\lim_{n\in A_j}\{x_n^i\} = a_j^i$ para $j \in \{1,\ldots,n\}$. Definiendo $\boldsymbol{a} = (a_1,\ldots,a_n)$ se tiene que $\lim_{n\in A_n}\{\boldsymbol{x}_n\} = \boldsymbol{a}$, lo que concluye la demostración. ■

4.5. Conjuntos abiertos y cerrados

El conjunto básico con el cual definimos la topología de un espacio normado es

$$B(\boldsymbol{x}_0,\varepsilon) = \{\boldsymbol{x} \in E : \|\boldsymbol{x} - \boldsymbol{x}_0\| < \varepsilon\}$$

que recibe el nombre de bola abierta de radio ε y centro $\boldsymbol{x}_0$.

Definición 4.7. Un punto $\boldsymbol{x} \in E$ es interior a C si existe $\varepsilon > 0$ tal que $B(\boldsymbol{x},\varepsilon) \subseteq C$. y un conjunto $A \subseteq E$ se dice abierto si:

$$\forall \boldsymbol{x} \in A, \exists \varepsilon > 0 \text{ tal que } B(\boldsymbol{x},\varepsilon) \subseteq A$$

Proposición 4.4. Un conjunto es abierto $\Leftrightarrow$ todos sus puntos son interiores.

Proposición 4.5. (Propiedades de los abiertos)

1. Si $A_1, A_2, \ldots, A_n$ son abiertos entonces $\bigcap_{i=1}^n A_i$ es abierto.
2. Si $\{A_i\}_{i\in I}$ es una familia de abiertos entonces $\bigcup_{i\in I} A_i$ es abierto.
3. E es abierto, $\varnothing$ es abierto.

Sea $\tau = \{A \subseteq E : A \text{ es abierto}\}$, τ se conoce como topología en E gracias a que satisface las propiedades 1,2 y 3 de los abiertos.

Definición 4.8. Sea $A \subseteq E$. Un punto $\boldsymbol{x} \in E$ se dice punto de acumulación de A si

$$\forall \varepsilon > 0 \; A \cap (B(\boldsymbol{x},\varepsilon) \backslash \{\boldsymbol{x}\}) \neq \varnothing$$

Nota 4.6. Para ser punto de acumulación de A no es necesario pertenecer a A. Por otra parte, no es suficiente estar en A para ser de acumulación.

Ejemplo 4.16.

1. Sea $A = (0,1)$. Entonces $x = 0$ es punto de acumulación de A.
2. Sea $A = (0,1) \cup \{3\}$. Luego $x = 3 \in A$, pero x no es punto de acumulación.

Definición 4.9. $A \subseteq E$ es cerrado si A contiene a todos sus puntos de acumulación.

Proposición 4.6. $A \subseteq \mathbb{R}^n$ es cerrado si y sólo si:

$$\text{Para toda sucesión } \{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subseteq A : (\,(\boldsymbol{x}_k \to \boldsymbol{x}) \Rightarrow \boldsymbol{x} \in A)$$

Demostración.
($\Rightarrow$): Supongamos que A es cerrado. Sea $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subseteq A$ una sucesión cualquiera tal que $\lim\limits_{\boldsymbol{x}_k\to\boldsymbol{x}} \boldsymbol{x}_k = \boldsymbol{x}$. Queremos demostrar que $\boldsymbol{x} \in A$. Supongamos que no, es decir, que $\boldsymbol{x} \in A^C$. Como A es cerrado, existe $\varepsilon > 0$ tal que $B(\boldsymbol{x}, \varepsilon) \subset A^C$. Por otra parte, de la definición de límite, existe $k_0 \in \mathbb{N}$ tal que $\boldsymbol{x}_k \in B(\boldsymbol{x}, \varepsilon)$ para todo $k \geq k_0$, lo que es imposible pues $\boldsymbol{x}_k \in A$.

($\Leftarrow$): Para demostrar la recíproca probemos la contrarrecíproca. Es decir, si A no es cerrado entonces existe una sucesión $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subset A$ que converge a $\boldsymbol{x}$ y $\boldsymbol{x} \notin A$.

Como A no es cerrado, A^C no es abierto, entonces existe un punto $\boldsymbol{x} \in A^C$ tal que

$$\text{Para todo } \varepsilon > 0 \text{ se tiene } B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing$$

Esta proposición nos permite construir una sucesión $\{\boldsymbol{x}_k\} \subset A$ de la siguiente manera: para cada $k \in \mathbb{N}$ tomamos $\varepsilon = \frac{1}{k}$ entonces, como $B(\boldsymbol{x}, 1/k) \cap A \neq \varnothing$, podemos elegir $\boldsymbol{x}_k \in B(\boldsymbol{x}, \frac{1}{k})$ y $\boldsymbol{x}_k \in A$. Por definición esta sucesión converge a $\boldsymbol{x}$, concluyendo la demostración pues $\boldsymbol{x} \notin A$. ■

Teorema 4.6. *$A \subseteq E$ es cerrado $\Leftrightarrow$ A^C es abierto.*

Demostración.
($\Rightarrow$): Supongamos que A es cerrado. Sea $\boldsymbol{x} \in A^C$, entonces $\boldsymbol{x}$ no es punto de acumulación de A, por lo tanto existe $\varepsilon > 0$ tal que

$$A \cap (B(\boldsymbol{x}, \varepsilon) \backslash \{\boldsymbol{x}\}) = \varnothing$$

pero esto es equivalente a $B(\boldsymbol{x}, \varepsilon) \backslash \{\boldsymbol{x}\} \subseteq A^C$ y como además $\boldsymbol{x} \in A^C$, tenemos que $B(\boldsymbol{x}, \varepsilon) \subseteq A^C$, y por lo tanto A^C es abierto.

($\Leftarrow$): Supongamos ahora que A^C es abierto. Sea $\boldsymbol{x} \in E$ punto de acumulación de A. Debemos demostrar que $\boldsymbol{x} \in A$. Por contradicción, si $\boldsymbol{x} \in A^C$, como es abierto,$\exists \varepsilon > 0$ talque

$$B(\boldsymbol{x}, \varepsilon) \subseteq A \Rightarrow B(\boldsymbol{x}, \varepsilon) \cap A = \varnothing \Rightarrow B(\boldsymbol{x}, \varepsilon) \backslash \{\boldsymbol{x}\} \cap A = \varnothing \quad \Rightarrow\Leftarrow$$

tenemos una contradicción, pues $\boldsymbol{x}$ es punto de acumulación de A. Por lo tanto $\boldsymbol{x} \in A$. ■

Ejemplo 4.17. $\mathbb{R}^n$ y el conjunto vacío $\varnothing$, son conjuntos abiertos. Aún cuando $\mathbb{R}^n$ es obviamente abierto, el caso del conjunto vacío requiere una reflexión. Si $\varnothing$ no es abierto entonces existe $\boldsymbol{x}_0 \in \varnothing$ para el cual $B(\boldsymbol{x}_0, r) \cap \mathbb{R}^n \neq \varnothing$, para cada $r > 0$. Esto es absurdo pues no hay elementos en $\varnothing$.

Ejemplo 4.18. El conjunto $A = \{(x, y) : x > 1\}$ es un conjunto abierto. En efecto, si $(x, y) \in A$ entonces $B((x, y), \frac{x-1}{2}) \subset A$.

Ejemplo 4.19. Si $x_0 \in \mathbb{R}$ y $r > 0$ entonces $B(x_0, r)$ es un conjunto abierto. En efecto, si $x \in B(x_0, r)$ entonces $B(x, (r - \|x - x_0\|)/2) \subset B(x_0, r)$. Usando la desigualdad triangular muestre la veracidad de esta última afirmación y haga un dibujo.

Definición 4.10. Diremos que $A \subseteq \mathbb{R}^n$ es un conjunto cerrado si A^C es abierto.

Nota 4.7. Los conjuntos $\mathbb{R}^n$ y $\varnothing$ son abiertos y cerrados. También notamos que hay conjuntos que no son abiertos ni cerrados. Ver ejemplo a continuación.

Ejemplo 4.20. Sean $A = \{\frac{1}{n} : n \in \mathbb{N}\}$ y $B = A \cup \{0\}$. Entonces:

1. A no es cerrado. En efecto A^C no es abierto, pues $0 \in A^C$ y $(\forall r > 0)$ $B(0, r) \not\subseteq A^C$. Cualquiera sea $r > 0$ se tiene que
$$(\exists n \in \mathbb{N}) \text{ tal que } \frac{1}{n} < r$$
Es decir, $\frac{1}{n} \in B(0, r)$.

2. A no es abierto pues $1 \in A$, pero $(\forall r > 0)$ $B(x_0, r) \not\subseteq A$.

3. B es cerrado y no es abierto.

Teorema 4.7. *Sean $\|\cdot\|_1$ y $\|\cdot\|_2$ dos normas equivalentes en el espacio E. $A \subseteq E$ es abierto en $(E, \|\cdot\|_1) \Leftrightarrow A$ es abierto en $(E, \|\cdot\|_2)$.*

Demostración.
$(\Rightarrow)$: Recordemos que existen constantes $c_1 > 0$, $c_2 > 0$ tales que
$$c_1\|\boldsymbol{x}\|_1 \leq \|\boldsymbol{x}\|_2 \leq c_2\|\boldsymbol{x}\|_1 \ \forall \boldsymbol{x} \in E$$
Supongamos que $A \subseteq E$ es abierto en $(E, \|\cdot\|_1)$.Sea $\boldsymbol{x} \in A$, entonces existe $\varepsilon > 0$ tal que
$$\{\boldsymbol{y} \in E : \|\boldsymbol{x} - \boldsymbol{y}\|_1 < \varepsilon\} \subseteq A$$
pero entonces
$$\{\boldsymbol{y} \in E : \|\boldsymbol{x} - \boldsymbol{y}\|_2 < c_1\varepsilon\} \subseteq \{\boldsymbol{y} \in E : \|\boldsymbol{x} - \boldsymbol{y}\|_1 < \varepsilon\} \subseteq A$$
o sea, encontramos $\varepsilon c_1 > 0$ tal que
$$B_{\|\cdot\|_2}(\boldsymbol{x}, \varepsilon c_1) \subseteq A$$
Así que todos los puntos de A son interiores con la norma $\|\cdot\|_2$ lo que implica que A es abierto en $(E, \|\cdot\|_2)$.

$(\Leftarrow)$: *Tarea.* ■

Nota 4.8. Todas las nociones que se definan a partir de los conjuntos abiertos de $(E, \|\cdot\|)$ quedan inalteradas cuando se cambia $\|\cdot\|$ por una norma equivalente.

Definición 4.11. Sea $A \subseteq E$. Definimos los siguientes conjuntos

1. Derivado de A:
$\text{der}(A) = \{\boldsymbol{x} \in E : \boldsymbol{x} \text{ es punto de acumulación de } A\}$

2. Adherencia o cerradura de A:
$\text{adh}(A) = A \cup \text{der}(A) = A \cup \{\boldsymbol{x} \in E : \boldsymbol{x} \text{ es punto de acumulación de } A\}$

3. Interior de A:
   $\text{int}(A) = \{\boldsymbol{x} \in A : \boldsymbol{x} \text{ es punto interior de } A\}$

4. Frontera de A:
   $\text{fr}(A) = \text{adh}(A) \setminus \text{int}(A)$

Nota 4.9. $\text{int}(A) \subset A$ y $A \subset \text{adh}(A)$. Observamos que un punto adherente de A no necesita estar en A, así mismo un punto de acumulación y un punto frontera no necesitan estar en A.

Se obtiene de las definiciones la siguiente proposición:

Proposición 4.7. *A es abierto si y sólo si $A = \text{int}(A)$ y A es cerrado si y sólo si $A = \text{adh}(A)$.*

Ejemplo 4.21. En $\mathbb{R}$ sea $A = [1, 2) \cup \{3\}$. Entonces: $\text{der}(A) = [1, 2]$, $\text{int}(A) = (1, 2)$, $\text{adh}(A) = [1, 2] \cup \{3\}$ y $\text{fr}(A) = \{1, 2, 3\}$.

Ejemplo 4.22. Demuestre que $\boldsymbol{x} \in \text{adh}(A)$ si y sólo si existe una sucesión $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}} \subset A$ tal que $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$.

Solución.
($\Rightarrow$) A es cerrado, entonces toda sucesión convergente en A tiene su límite en A.
Sea la sucesión $\{\boldsymbol{x}_k\}, k \in \mathbb{N} : \boldsymbol{x}_k \to \boldsymbol{x} \Rightarrow \boldsymbol{x} \in A$

$$\boldsymbol{x}_k \to \boldsymbol{x} \Leftrightarrow \forall \varepsilon > 0 \exists k_0 \in \mathbb{N} : \forall k \geq k_0, \boldsymbol{x}_k \in B(\boldsymbol{x}, \varepsilon)$$

Pero $\boldsymbol{x}_k \in A$, entonces $\forall \varepsilon > 0$ $B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing \Rightarrow \boldsymbol{x} \in \text{adh}(A)$. Sabemos que A es cerrado, lo que en otras palabras es $A = \text{adh}(A)$ lo cual implica que $\boldsymbol{x} \in A$.

($\Leftarrow$) Toda sucesión convergente del conjunto A tiene su límite en el conjunto, entonces A es cerrado. Supongamos que $\boldsymbol{x} \in \text{adh}(A)$, entonces se tiene que

$$B(\boldsymbol{x}, \varepsilon) \cap A \neq \varnothing \ \forall \varepsilon > 0$$

Escojamos $\varepsilon = 1/k$, entonces debe tenerse

$$B(\boldsymbol{x}, 1/k) \cap A \neq \varnothing \ \forall k \in \mathbb{N}$$

lo que implica que existe $\boldsymbol{x}_k$ perteneciente a $B(\boldsymbol{x}, 1/k) \cap A$ para todo $k \in \mathbb{N}$. Entonces, $\{\boldsymbol{x}_k\}$ está en A y $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$ □

4.6. Conjuntos compactos

Motivación: Podemos pensar en los conjuntos compactos como aquellos que tienen un radio finito o mejor dicho cuyas dimensiones son acotadas. En diversos problemas de ingeniería es común encontrarse con problemas de maximización o minimización. Si maximizamos o minimizamos determinada función sobre un compacto gracias a algunos teoremas, que veremos en el capítulo siguiente, nos aseguramos de que la solución existe.

Definición 4.12. (Conjunto compacto)
En el caso de conjuntos de $\mathbb{R}^n$, como ya veremos a continuación, podemos decir que $C \subset \mathbb{R}^n$ es compacto si es cerrado y acotado a la vez.

De forma general un conjunto $C \subset \mathbb{R}^n$ es compacto si para toda familia de conjuntos abiertos $\{A_i\}_{i\in\Lambda}$ tal que $C \subset \bigcup_{i\in\Lambda} A_i$, existe un subconjunto finito $\Lambda^* \subset \Lambda$ tal que $C \subset \bigcup_{i\in\Lambda^*} A_i$. Esto es, un conjunto C es compacto cuando toda cobertura por abiertos de C tiene una subcobertura finita.

Si suponemos que un conjunto C^* es compacto se demuestra por contradicción que no lo es si encontramos una cobertura por abiertos de C^* que no tenga una subcobertura finita.

Ejemplo 4.23. Para fijar ideas, consideremos los siguientes conjuntos:

1. $(0,1)$ en los reales no es compacto, ya que la cobertura por abiertos $\{(\frac{1}{n},1), n \in \mathbb{N}\}$ no tiene una subcobertura finita. Notemos que $(0,1)$ tiene una cobertura por abiertos definida por $\bigcup_{i\in\Lambda} A_i = (\frac{1}{2},1) \cup (\frac{1}{3},1) \cup \ldots \cup (\frac{1}{n},1)$. Sin embargo, no existe una subcobertura finita por abiertos que cubra completamente a $(0,1)$. Esto último es porque el ínfimo de cualquier subcobertura finita puede acercarse a cero pero, en la medida que la cantidad de elementos de la subcobertura sea finita, el valor del ínfimo queda "lejos" de cero y entonces es posible converger y acercarse lo suficiente a cero en la medida que la cantidad de elementos de la subcobertura tienda a infinito.

2. $[0,1)$ en los reales no es compacto, ya que la cobertura por abiertos $\{(-1,1-\frac{1}{n}), n \in \mathbb{N}\}$ no tiene una subcobertura finita. En este caso el extremo derecho no es cerrado por lo que se puede construir una sucesión que converge a uno y este hecho impide que el conjunto tenga una cobertura finita.

3. $[0,+\infty)$ en los reales no es compacto, ya que $\{(-1,n), n \in \mathbb{N}\}$ no tiene una subcobertura finita. En este caso se puede construir una sucesión que no converge y tiende a infinito lo cual impide que el conjunto tenga una cobertura finita.

4. $[0,1]$ en los racionales no es compacto pero si lo es en los reales. En el caso de los racionales $\{[\frac{1}{n},1], n \in \mathbb{Q}\}$ no tiene una subcobertura finita.
   Este caso debe ser analizado aparte, pues la cobertura abarca completamente el conjunto $[0,1]$ en los racionales pero no tiene subcobertura finito alguno. Para el caso en que $[0,1]$ se define en los reales se puede razonar por contradicción: Supongamos que $[0,1]$ tiene una cobertura por abiertos $\bigcup_{i\in\Lambda} A_i$ pero no existe un subcobertura finita por abiertos $\bigcup_{i\in\Lambda^*} A_i$ de $\bigcup_{i\in\Lambda} A_i$. Entonces $[0,\frac{1}{2}]$ y $[\frac{1}{2},0]$ no pueden ser cubiertos por $\bigcup_{i\in\Lambda^*} A_i$. Digamos, de manera arbitraria, que $[a_1,b_1]$ es cualquiera de estos dos intervalos y entonces $[a_1,\frac{1}{2}(b_1-a_1)]$ y $[\frac{1}{2}(b_1-a_1),b_1]$ tampoco pueden ser cubiertos por $\bigcup_{i\in\Lambda^*} A_i$. Digamos que $[a_2,b_2]$ es cualquiera de estos dos intervalos y así mediante un razonamiento inductivo se pueden definir dos sucesiones $\{a_i\}_{i=1}^n$ y $\{b_i\}_{i=1}^n$ en $[0,1]$ tales que
   $$a_n \leq a_{n+1} < b_{n+1} \leq b_n$$
   $$b_n - a_n = \frac{1}{2^n}$$
   $$\bigcup_{i\in\Lambda^*} A_i \subset [a_n,b_n]$$
   Las primeras dos ecuaciones nos dicen que, intuitivamente, es posible encontrar un valor $c \in [0,1]$ tal que $c = \bigcap_{i=1}^{\infty}[a_i,b_i]$ y la tercera ecuación nos dice que $[a_n,b_n]$ no puede ser cubierto por $\bigcup_{i\in\Lambda^*} A_i$. Supongamos que $c \in \bigcup_{i\in\Lambda^*} A_i$, entonces c está contenido en un abierto y además se tiene que $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = c$, entonces $[a_n,b_n] \in \bigcup_{i\in\Lambda^*} A_i$ en la medida que $n \to \infty$ y esto nos dice que efectivamente $[a_n,b_n]$ es cubierto por $\bigcup_{i\in\Lambda^*} A_i$. Llegamos a una contradicción y por definición $[0,1]$ es compacto, luego es posible concluir que toda cobertura por abiertos del conjunto tiene una subcobertura finita.

Definición 4.13. (Conjunto secuencialmente compacto)
Sea $C \subset \mathbb{R}^n$, diremos que este conjunto es secuencialmente compacto si toda sucesión convergente en C tiene una subsucesión convergente en C. La misma definición aplica si reemplazamos $\mathbb{R}^n$ por un e.v.n. E.

Teorema 4.8. *Todo conjunto $S \subset C$ cerrado es compacto si $C \subset \mathbb{R}^n$ es compacto.*

Demostración. Sea S cerrado y $\{A_i\}_{i\in\Lambda}$ una cobertura por abiertos de S, entonces $\{A_i\}_{i\in\Lambda} \cup S^C$ es una cobertura por abiertos de C. Como C es compacto, tiene una subcobertura finita por abiertos $\{A_i\}_{i\in\Lambda^*} \cup S^C$ que además es una subcobertura finita por abiertos de $\{A_i\}_{i\in\Lambda} \cup S^C$. Entonces, $\{A_i\}_{i\in\Lambda^*}$ es una subcobertura finita por abiertos de $\{A_i\}_{i\in\Lambda}$ lo que concluye la demostración. ■

Teorema 4.9. *Si $C \subset \mathbb{R}^n$ es compacto, entonces C es cerrado y acotado.*

Demostración. Para demostrar que es acotado, partamos de la base que C es compacto y diferente de vacío. Por definición, sea $\{A_i\}_{i\in\Lambda}$ una cobertura por abiertos de C, entonces necesariamente esta cobertura tiene un subcobertura finita $\{A_i\}_{i\in\Lambda^*}$ tal que $C \subset \bigcup_{i\in\Lambda^*} A_i$. Sea $n = \text{máx}_{i\in\Lambda^*} i$, entonces $C \subset \bigcup_{i=1}^{n} A_i$, como Λ^* tiene máximo, C está contenido completamente en una unión finita lo que demuestra que es acotado. El razonamiento no es válido si C no es una parte finita de $\mathbb{R}^n$ pero ese no es el caso que debemos demostrar.

Ahora debemos demostrar que C es cerrado. Los casos $C = \mathbb{R}^n$ y $C = \varnothing$ significan que no hay nada que demostrar. Supongamos que $C \cap \mathbb{R}^n = C$, entonces escojamos $\boldsymbol{x} \in C$ e $\boldsymbol{y} \in C^C$. Se tendrá que dado $\varepsilon > 0$ existe $B(\boldsymbol{x}, \varepsilon)$ tal que $B(\boldsymbol{x}, \varepsilon) \cap B(\boldsymbol{y}, \varepsilon) = \varnothing$, por ejemplo, si escogemos $\varepsilon = \frac{1}{2} d(\boldsymbol{x}, \boldsymbol{y})$. Sea $\{A_i\}_{i\in\Lambda}$ una cobertura por abiertos de C, debe existir $\{A_i\}_{i\in\Lambda^*} \subset \{A_i\}_{i\in\Lambda}$ que corresponde a una subcobertura finita por abiertos de C. Definamos $\varepsilon^* = \text{mín}_{\boldsymbol{x}\in\{A_i\}} d(\boldsymbol{x}, \boldsymbol{y}) : i \in \Lambda^*$ y entonces $B(\boldsymbol{y}, \varepsilon^*) \cap C^C \neq \varnothing$ y $B(\boldsymbol{y}, \varepsilon^*) \subset C^C$ por lo que C^C es abierto, lo que concluye la demostración. ■

Teorema 4.10. (Heine-Borel)
Sea $C \subset \mathbb{R}^n$. C es compacto si y sólo si es un conjunto cerrado y acotado.

Demostración.
($\Rightarrow$): Esta implicancia ya fue demostrada en el teorema 4.9.

($\Leftarrow$): Si C es acotado, existe $\varepsilon > 0$ tal que $C \subset B(\boldsymbol{x}, \varepsilon)$ para $\boldsymbol{x} \in C$. Entonces, C corresponde a un subconjunto cerrado del rectángulo $[a, b] \times \ldots \times [a, b]$. Escojamos $a = \text{mín}\{x_i - \varepsilon, i = 1, \ldots, n\}$ y $b = \text{máx}\{x_i + \varepsilon, i = 1, \ldots, n\}$, con lo cual $[a, b]^n$ es compacto y el teorema 4.8 permite concluir que C es compacto lo cual finaliza la demostración. ■

Nota 4.10. Establecer que un conjunto compacto debe ser cerrado y acotado y viceversa nos da una relación de si y sólo si que puede extenderse a todo espacio de dimensión finita, pero en espacios de dimensión infinita es falsa. De hecho, la demostración del teorema de Heine-Borel que hicimos sólo es válida con $n \in \mathbb{N}$, es decir $1 \leq n < \infty$.

Proposición 4.8. En $\mathbb{R}^n$ las siguientes propiedades son equivalentes:

1. $C \subset \mathbb{R}^n$ es compacto.
2. $C \subset \mathbb{R}^n$ es cerrado y acotado.
3. Toda sucesión en $C \subset \mathbb{R}^n$ tiene una subsucesión convergente.

4.7. Consecuencias de la compacidad

Daremos algunas consecuencias no menores que resultan de los conjuntos compactos. Su importancia radica en la existencia de elementos que maximizan o minimizan funciones cuyo dominio es compacto.

Teorema 4.11. *Sea $f : U \subset \mathbb{R}^n \to \mathbb{R}$ una función continua. Si $C \subset U$ es compacto, entonces $f(C)$ es compacto.*

Demostración. Escojamos arbitrariamente cualquier familia de conjuntos abiertos tales que $f(C) \subset \bigcup_{i\in\Lambda} A_i$, es decir estamos fijando una cobertura por abiertos $\{A_i\}_{i\in\Lambda}$ de $f(C)$. Se tendrá que $C \subset f^{-1}(\bigcup_{i\in\Lambda} A_i) = \bigcup_{i\in\Lambda} f^{-1}(A_i)$, lo cual significa que existe una subcobertura finita $\bigcup_{i\in\Lambda^*} f^{-1}(A_i)$ de C, con $\Lambda^* \subset \Lambda$. Entonces, $f(C)$ es compacto ya que $f(C)$ puede ser cubierto por una cantidad finita de elementos de $\{A_i\}_{i\in\Lambda}$. ■

Definición 4.14. Sea $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ una función no necesariamente continua. Diremos que f alcanza su máximo (respectivamente mínimo) en U, si existe $\boldsymbol{x}_M \in U$ (respectivamente $\boldsymbol{x}_m \in U$) tal que $f(\boldsymbol{x}) \leq f(\boldsymbol{x}_M)$ (respectivamente $f(\boldsymbol{x}_m) \leq f(\boldsymbol{x})$), para todo $\boldsymbol{x} \in U$.

De esto se obtiene un corolario importante:

Corolario 4.1. *Toda función continua $f : C \subset \mathbb{R}^n \to \mathbb{R}$ definida en un conjunto compacto C alcanza su máximo y su mínimo en C.*

Demostración. Como f es continua, $f(C)$ es compacto. De esta forma, $f(C)$ es cerrado y acotado. Por ser $f(C)$ acotado, existe $\boldsymbol{x}^i = \inf\{\boldsymbol{x} : \boldsymbol{x} \in f(C)\}$ y $\boldsymbol{x}^s = \sup\{\boldsymbol{x} : \boldsymbol{x} \in f(C)\}$. Por definición de ínfimo y supremo, podemos construir las sucesiones $\{\boldsymbol{x}_n^i\}_{n\in\mathbb{N}} \subset f(C)$ y $\{\boldsymbol{x}_n^s\}_{n\in\mathbb{N}} \subset f(C)$ que convergen, respectivamente, a $\boldsymbol{x}^i$ y $\boldsymbol{x}^s$. Por ser $f(C)$ cerrado, $\boldsymbol{x}^i$ y $\boldsymbol{x}^s$ están en $f(C)$, lo cual concluye la demostración. ■

A partir del corolario 4.1 se tiene la siguiente consecuencia:

Corolario 4.2. *En $\mathbb{R}^n$ todas las normas son equivalentes.*

Demostración. La demostración ya fue hecha para el teorema 4.2. El corolario proviene de que las normas definen funciones continuas (la demostración de esto queda de *tarea*). Definiendo $S = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| = 1\}$ se define un conjunto cerrado y acotado en $(\mathbb{R}^n, \|\cdot\|)$ para cualquier norma. Sobre este hecho, los teoremas 4.9 y 4.1 permiten concluir la equivalencia. ■

Definición 4.15. Sean $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ espacios vectoriales normados y $f : A \subseteq E \to F$. f es uniformemente continua en A si

$$(\forall\varepsilon > 0)(\exists\delta > 0)(\forall \boldsymbol{x}, \boldsymbol{y} \in A) \text{ tal que } \|\boldsymbol{x} - \boldsymbol{y}\|_E \leq \delta \Rightarrow \|f(\boldsymbol{x}) - f(\boldsymbol{y})\|_F \leq \varepsilon$$

Nota 4.11. En el caso de la continuidad uniforme cambia el orden de los cuantificadores. La razón de esto es que en la continuidad uniforme el valor de δ no depende del $\boldsymbol{x}$ que se escoja.

Proposición 4.9. Si f es uniformemente continua en A, entonces f es continua en A. La recíproca es, en general, falsa. Un claro ejemplo de esto último es que la función $f(x) = x^2$ es continua pero no es uniformemente continua en $\mathbb{R}$.

Finalizamos esta sección con el siguiente teorema:

Teorema 4.12. *Sean $(E, \|\cdot\|_E)$ y $(F, \|\cdot\|_F)$ espacios vectoriales normados. Sea $f : A \subseteq E \to F$ continua y A compacto. Entonces f es uniformemente continua en A.*

Demostración. Supongamos que f no es uniformemente continua en A. Entonces

$$\exists\varepsilon > 0\, \forall\delta > 0\, \exists \boldsymbol{x}, \boldsymbol{y} \in A \text{ tal que } \|\boldsymbol{x} - \boldsymbol{y}\|_E < \delta \,\wedge\, \|f(\boldsymbol{x}) - f(\boldsymbol{y})\|_F > \varepsilon$$

Sea $n \in \mathbb{N}$, tomamos $\delta = 1/n > 0$ y escogemos $\boldsymbol{x}_n, \boldsymbol{y}_n \in A$ tales que

$$\|\boldsymbol{x}_n - \boldsymbol{y}_n\|_E < \frac{1}{n} \ \wedge \ \|f(\boldsymbol{x}_n) - f(\boldsymbol{y}_n)\|_F > \varepsilon$$

Como $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}} \subseteq A$ y A es compacto, existe una subsucesión $\{x_{n_k}\}_{n\in\mathbb{N}}$ de $\{x_n\}_{n\in\mathbb{N}}$ que converge a un $\boldsymbol{x} \in A$. A su vez la sucesión $\{\boldsymbol{y}_{n_k}\}_{k\in\mathbb{N}} \subseteq A$ posee una subsucesión $\{\boldsymbol{y}_{n_{k_l}}\}_{l\in\mathbb{N}}$ convergente, $\boldsymbol{y}_{n_{k_l}} \to \boldsymbol{y} \in A$. Notemos que $\{\boldsymbol{x}_{n_{k_l}}\}$ es una subsucesión de $\{\boldsymbol{x}_{n_k}\}$ y por lo tanto también converge a $\boldsymbol{x}$. Lo anterior implica que $(\boldsymbol{x}_{n_{k_l}} - \boldsymbol{y}_{n_{k_l}}) \to (\boldsymbol{x} - \boldsymbol{y})$ cuando $l \to \infty$, pero por otra parte tenemos que

$$\|\boldsymbol{x}_n - \boldsymbol{y}_n\|_E < \frac{1}{n} \,\forall n \in \mathbb{N} \ \Rightarrow \ \|\boldsymbol{x}_{n_{k_l}} - \boldsymbol{y}_{n_{k_l}}\|_E < \frac{1}{n_{k_l}} \to 0$$

es decir, $(\boldsymbol{x}_{n_{k_l}} - \boldsymbol{y}_{n_{k_l}}) \to 0$ cuando $l \to \infty$ lo que implica que $\boldsymbol{x} = \boldsymbol{y}$.
Por la elección de las sucesiones tenemos que

$$\|f(\boldsymbol{x}_{n_{k_l}}) - f(\boldsymbol{y}_{n_{k_l}})\|_F > \varepsilon \ \forall l \in \mathbb{N}$$

Tomando límite cuando $l \to \infty$ y gracias a la continuidad de f obtenemos que

$$\|f(\boldsymbol{x}) - f(\boldsymbol{y})\|_F \geq \varepsilon > 0$$

lo que es imposible pues $\boldsymbol{x} = \boldsymbol{y}$. ■

4.8. Conjuntos convexos

Motivación: La importancia de los convexos es que estos se pueden separar matemáticamente y sus propiedades de separación nos sirven para demostrar que la solución de un problema es la mejor de todas las soluciones posibles. Retomaremos esta idea en los capítulos siguientes.

Definición 4.16. Un conjunto C es convexo si el segmento que une dos puntos pertenecientes al conjunto también pertenece a C.

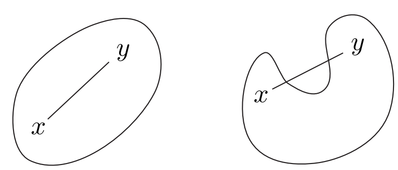


Figura 4.2: Conjunto convexo y no convexo respectivamente

De manera formal, un conjunto C es convexo si dados dos elementos $\boldsymbol{x}, \boldsymbol{y} \in C$ y un número real $\lambda \in [0, 1]$ se cumple

$$\boldsymbol{z} \in C \Leftrightarrow \boldsymbol{z} = \lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \tag{4.4}$$

Es decir, (4.4) es una combinación lineal que define un segmento de extremos $\boldsymbol{x}, \boldsymbol{y}$.

Definición 4.17. Sean $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^n$ y $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que $\sum_{i=1}^n \lambda_i = 1$. El vector $\boldsymbol{y} = \sum_{i=1}^n \lambda_i \boldsymbol{x}_i$ es una combinación convexa de $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

Proposición 4.10. Todo espacio vectorial define un conjunto convexo.

Demostración. No es necesario demostrar. Como todo espacio vectorial, por definición, es cerrado para la suma y la ponderación por escalar se tiene que es convexo. ■

Ahora que ya tenemos la definición de convexidad podemos agregar algo más sobre las normas. Ya habíamos señalado que en el caso de la norma ρ se debe cumplir que $\rho \geq 1$ porque además de la desigualdad triangular se debe cumplir que el conjunto

$$C = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\|_\rho \leq \alpha,\ \alpha \in \mathbb{R}\}$$

debe ser convexo. Con $\rho < 1$ se tiene que C no es convexo y el ejemplo 4.2 es útil para fijar ideas.

Proposición 4.11. Sea $C \subset \mathbb{R}^n$ convexo. Entonces $\text{int}(C)$ y $\text{adh}(C)$ son convexos.

Demostración. Sean $\boldsymbol{x}, \boldsymbol{y} \in \text{int}(C)$ y $\boldsymbol{z} = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$. Dado que $\boldsymbol{x}, \boldsymbol{y}$ son interiores a C, existe $r > 0$ tal que $B(\boldsymbol{x}, r),\ B(\boldsymbol{y}, r) \in C$.

Sea $v \in \mathbb{R}^n$ y $\|\boldsymbol{v}\| < r$, entonces $\boldsymbol{x} + \boldsymbol{v} \in B(\boldsymbol{x}, r)$ y además $\boldsymbol{y} + \boldsymbol{v} \in B(\boldsymbol{y}, r)$. Luego

$$\begin{aligned} \boldsymbol{z} + \boldsymbol{v} &= (\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + (\lambda\boldsymbol{v} + (1-\lambda)\boldsymbol{v}) \\ &= \lambda(\boldsymbol{x} + \boldsymbol{v}) + (1-\lambda)(\boldsymbol{y} + \boldsymbol{v}) \end{aligned}$$

Tenemos que $\boldsymbol{z} + \boldsymbol{v} \in C$ dado que C es convexo. Definamos $\boldsymbol{v} = \boldsymbol{z}_0 - \boldsymbol{z}$ tal que $\boldsymbol{z}_0 \in B(\boldsymbol{z}, r)$, entonces $\|v\| < r$ y se obtiene $\boldsymbol{z}_0 = \boldsymbol{z} + \boldsymbol{v} \in C$. Por lo tanto $B(\boldsymbol{x}, r) \subset C$, $\boldsymbol{z} \in \text{int}(C)$ y llegamos a que $\text{int}(C)$ es convexo.

Supongamos ahora que $\boldsymbol{x}, \boldsymbol{y} \in \text{adh}(C)$. Sean $\boldsymbol{z} = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ y $r_0 > 0$, entonces existen $\boldsymbol{v}, \boldsymbol{w}$ tales que $\|\boldsymbol{v}\| < r_0$, $\|\boldsymbol{w}\| < r_0$, $\boldsymbol{x} + \boldsymbol{v} \in C$ y además $\boldsymbol{y} + \boldsymbol{w} \in C$. Llegamos a que $\boldsymbol{z}_0 = \lambda(\boldsymbol{x} + \boldsymbol{v}) + (1-\lambda)(\boldsymbol{y} + \boldsymbol{w}) \in C$ y por lo tanto $\|\boldsymbol{z}_0 - \boldsymbol{z}\| \leq \lambda\|\boldsymbol{v}\| + (1-\lambda)\|\boldsymbol{w}\| < r_0$, $\boldsymbol{z} \in \text{adh}(C)$ que permite concluir que $\text{adh}(C)$ es un conjunto convexo. ■

4.9. Ejercicios

Base algebraica y geométrica de $\mathbb{R}^n$

Ejercicio 49. Sean $\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_{n-1} \in \mathbb{R}^n$ tales que $\boldsymbol{x}_1 - \boldsymbol{x}_0, ..., \boldsymbol{x}_{n-1} - \boldsymbol{x}_0$ son linealmente independientes. Probar que existe exactamente un hiperplano conteniendo a $\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_{n-1}$.

Ejercicio 50. Sea $\boldsymbol{x}$ un vector cualquiera de $\mathbb{R}^n$ y $\boldsymbol{d}$ es un vector unitario:

1. Demuestre que $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, donde $\boldsymbol{y}$ es un múltiplo de d y $\boldsymbol{z}$ es perpendicular a $\boldsymbol{d}$.
2. Demuestre que los vectores $\boldsymbol{y}$ y $\boldsymbol{z}$ de la parte 1. están determinados unívocamente.

Ejercicio 51. Sea $A \in M_{n\times n}(\mathbb{R})$ una matriz simétrica.

1. Pruebe que existen vectores $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \in \mathbb{R}^n, \lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tales que:
$$A\boldsymbol{x} = \sum_{i=1}^{n} \lambda_i (\boldsymbol{x} \cdot \boldsymbol{e}_i)\boldsymbol{e}_i \qquad \forall \boldsymbol{x} \in \mathbb{R}^n.$$
Indicación. Piense en los valores y vectores propios de A.
2. Pruebe usando lo anterior que $\|A\boldsymbol{x}\| \leq C\|\boldsymbol{x}\|$.

Normas y conjuntos en un e.v.n

Ejercicio 1. Dados dos conjuntos A, B en un espacio vectorial normado E, demuestre

1. $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$
2. $\text{int}(A) \cup \text{int}(B) \subset \text{int}(A \cup B)$ (dé un ejemplo donde no hay igualdad)
3. $\text{adh}(A \cup B)) = \text{adh}(A) \cup \text{adh}(B)$.
4. $\text{adh}(A \cap B)) \subset \text{adh}(A) \cap \text{adh}(B)$ (dé un ejemplo donde no hay igualdad).
5. $\text{adh}(A) = \text{int}(A) \cup \text{fr}(A)$
6. $A \subset B \Rightarrow \text{int}(A) \subset \text{int}(B)$ y $\text{adh}(A) \subset \text{adh}(B)$
7. $\text{int}(A) \cap B = \varnothing \Leftrightarrow \text{int}(A) \cap \text{adh}(B) = \varnothing$
8. $\text{adh}(A) = E$ y $\text{int}(B) \cap A = \varnothing \Rightarrow \text{int}(B) = \varnothing$
9. $\text{int}(A^C) = (\text{adh}(A))^C$
10. $(\text{adh}(A^C)) = (\text{int}(A))^C$

Ejercicio 2. Sea
$$A = \{(x, y) \in \mathbb{R}^2 : (x-1)^2 + y^2 = \frac{1}{4} \wedge x < 1\}$$
Determine $\text{int}(A)$, $\text{adh}(A)$, $\text{fr}(A)$ y deduzca si A es un conjunto abierto o cerrado.

Ejercicio 3. Sea E el espacio vectorial de los polinomios de una variable real con coeficientes reales. Definamos la función $\|\cdot\|: E \to \mathbb{R}$ por

$$\|p\| = \max_{x\in[0,1]} |p(x)| \quad \forall p \in E$$

1. Demuestre que $\|\cdot\|$ es una norma en el espacio E.
   Hint: Use que un polinomio tiene exactamente tantos ceros como el grado, excepto si es el polinomio nulo.

2. Considere la función $\ell_1 : E \to \mathbb{R}$ definida por

   $$\ell_1(p) = p(1/2) \text{ para todo } p \in E$$

   Demuestre que:

3. la función ℓ_1 es lineal y verifica la desigualdad:

   $$\exists K \in \mathbb{R} \text{ tal que } |\ell_1(p)| \leq K\|p\| \text{ para todo } p \in E$$

4. La función ℓ_1 es continua.

Ejercicio 4. Sea $A \in \mathcal{M}_{n\times n}$ una matriz invertible. Se define $n : \mathbb{R}^n \to \mathbb{R}$ como

$$n = \|\boldsymbol{x}\| + \|A\boldsymbol{x}\|$$

Demuestre que n es una norma

Ejercicio 5. Dada una norma $\|\cdot\|$ en $\mathbb{R}^n$ y una matriz A de $n \times n$ invertible. Demuestre que la función $\|\boldsymbol{x}\|^* = \|A\boldsymbol{x}\|$ es una norma en $\mathbb{R}^n$.

Conjuntos abiertos y cerrados

Ejercicio 6. Analice el interior, la adherencia, el derivado y la frontera de

$$A = \bigcup_{k=2}^{\infty} B\left(\left(\frac{3}{2^{k+1}}, 0\right), \frac{1}{2^{k+1}}\right) \bigcup \overline{B}\left(\left(\frac{3}{4}, 0\right), \frac{1}{4}\right).$$

Ejercicio 7. Demuestre que:

1. $\mathrm{adh}(A)$ es un conjunto cerrado.

2. $\mathrm{int}(A)$ es un conjunto abierto.

3. $\mathrm{adh}(A)$ es el cerrado más pequeño que contiene a A y a su vez $\mathrm{int}(A)$ es el abierto más grande contenido en A.

4. $\mathbb{R}^n \setminus \mathrm{int}(A) = \mathrm{adh}(\mathbb{R}^n \setminus A)$.

Ejercicio 8. Pruebe que en general no se tiene que $\mathrm{int}(\mathrm{adh}(A)) = A$. Para esto siga los siguientes pasos:

1. Pruebe que $\mathrm{adh}(\mathbb{Q}) = \mathbb{R}$.

2. Pruebe que $\text{int}(\mathbb{R}) = \mathbb{R}$

3. Concluya.

Ejercicio 9. Sean $A, B \subset \mathbb{R}^n$. Se define $A + B = \{\boldsymbol{x} : \boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b}, \boldsymbol{a} \in A, \boldsymbol{b} \in B\}$.

1. Demuestre que $A + B$ es abierto si A es abierto

2. Dé un ejemplo en $\mathbb{R}$ donde A y B sean cerrados pero $A + B$ no sea cerrado.

Ejercicio 10. Verificar que el conjunto

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + xy < x\} \text{ es abierto en } \mathbb{R}^2$$

Ejercicio 11. Encontrar en $\mathbb{R}^2$ un conjunto que no sea abierto ni cerrado.

Ejercicio 12. Sea d la distancia en $\mathbb{R}^n$, asociada a alguna norma en $\mathbb{R}^n$ ($d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$). Sean $A \subseteq \mathbb{R}^n$ y $B \subseteq \mathbb{R}^n$ dos conjuntos no vacíos, demuestre:

1. $C = \{\boldsymbol{x} \in \mathbb{R}^n : d(\boldsymbol{x}, A) = d(\boldsymbol{x}, B)\}$ es cerrado

2. $D = \{\boldsymbol{x} \in \mathbb{R}^n : d(\boldsymbol{x}, A) < d(\boldsymbol{x}, B)\}$ es abierto donde $d(\boldsymbol{x}, A) = \text{ínf}\{d(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{y} \in A\}$.

3. Si A y B son cerrados y disjuntos entonces existen dos abiertos no vacíos U y V tales que $A \subset U$ y $B \subset V$ y $U \cap V = \varnothing$.

Ejercicio 13. Dada una matriz A de 2×2 considere su determinante y demuestre que el conjunto

$$B = \{A \in \mathcal{M}_2(\mathbb{R}) : A \text{ es invertible}\}$$

es abierto.

Indicación. Identifique $\mathcal{M}_2(\mathbb{R})$ con $\mathbb{R}^4$.

Sucesiones en un e.v.n

Ejercicio 14. Considere el espacio vectorial $C([0, 1], \mathbb{R})$ de las funciones continuas $f : [0, 1] \to \mathbb{R}$. Definamos la sucesión $\{f_k\}$ por:

$$f_k = \begin{cases} 1 & \text{si } x \in [0, 1/2] \\ -2^k(x - 1/2) + 1 & \text{si } x \in [1/2, 2^{-k} + 1/2] \\ 0 & \text{si } x \in [2^{-k} + 1/2, 1] \end{cases}$$

1. Verificar que $f_k \in C([0, 1], \mathbb{R})\ \forall k \in \mathbb{N}$.

2. Se define la norma $\|\cdot\|_1$ en $C([0, 1], \mathbb{R})$, como $\|f\|_1 = \int\limits_0^1 |f(x)|dx$. Demuestre que para esta norma $\{f_k\}$ es de Cauchy y no es convergente.

3. Sea $A([0, 1], \mathbb{R})$ el espacio vectorial de las funciones acotadas de $[0, 1]$ en $\mathbb{R}$, dotado de la norma $\|\cdot\|_\infty$, definida por $\|f\|_\infty = \sup\limits_{x \in [0,1]} |f(x)|$, demuestre que $\{f_k\}$ es convergente en este espacio.

Ejercicio 15. Sea $\|\cdot\|_1$ una norma en el e.v. E y $\{\boldsymbol{a}_k\}$ una sucesión de Cauchy respecto de dicha norma. Demostrar que si $\|\cdot\|_2$ es otra norma en E equivalente a $\|\cdot\|_1$, entonces $\{\boldsymbol{a}_k\}$ es sucesión de Cauchy también respecto a $\|\cdot\|_2$.
Si E es Banach respecto a $\|\cdot\|_1$. ¿Se puede decir que es Banach con $\|\cdot\|_2$?

Ejercicio 16. Demuestre que dos sucesiones de Cauchy $(\boldsymbol{x}_k)$ y $(\boldsymbol{y}_k)$ en $\mathbb{R}^n$ tienen el mismo límite sí y sólo sí $\lim\limits_{k\to\infty} \|\boldsymbol{x}_k - \boldsymbol{y}_k\| = 0$

Ejercicio 17. Demuestre la siguiente caracterización de espacios de Banach. Sea E un e.v.n.

$$E \text{ es un espacio de Banach}$$

sí y sólo sí

$$\left(\forall \{\boldsymbol{x}_n\}_{n\in\mathbb{N}} \subseteq E \quad \sum_{i=1}^{\infty} \|x_i\| < \infty\right) \Rightarrow \sum_{i=1}^{\infty} x_i \text{ converge en } E$$

Indicación. Para la implicación ($\Leftarrow$), dada una sucesión de Cauchy $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ en E, construya una subsucesión $\{\boldsymbol{x}_{n_k}\}_{n\in\mathbb{N}}$ tal que $\sum\limits_{k=1}^{\infty} \|\boldsymbol{x}_{n_{k+1}} - \boldsymbol{x}_{n_k}\| < \infty$. Luego concluya.

Ejercicio 18. Demuestre que el espacio $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ de todos las aplicaciones lineales ℓ de $\mathbb{R}^n$ en $\mathbb{R}^m$ es un e.v.n. en el que toda sucesión de Cauchy converge debido a que $\mathbb{R}^n$ es e.v.n. y en $\mathbb{R}^m$ toda sucesión de Cauchy converge.

Indicación. Primero demuestre que una aplicación lineal acotada es uniformemente continua y que si una aplicación lineal es continua en un punto, es acotada.

Compacidad

Ejercicio 19. Considere la sucesión $\{p_k\}$ en E definida por $p_k(\boldsymbol{x}) = \boldsymbol{x}^k$ y demuestre que

1. $\|p_k\| = 1$ para todo $k \in \mathbb{N}$.
2. $\{p_k\}$ no tiene punto de acumulación.
3. $B(0,1)$ en E no es compacta.

Ejercicio 20. Demuestre que si A es un conjunto compacto y $B \subset A$, entonces $\overline{B} = \operatorname{adh}(B)$ es un conjunto compacto.

Ejercicio 21.

1. Si $\{A_i\}_{i\in\mathbb{N}}$ es una familia decreciente de conjuntos compactos no vacíos en un e.v.n, demuestre que $\bigcap\limits_{i\in\mathbb{N}} A_i \neq \varnothing$.
2. Dé un ejemplo en $\mathbb{R}$ de una familia decreciente de conjuntos acotados (no vacíos) cuya intersección sea vacía.
3. Dé un ejemplo en $\mathbb{R}$ de una familia decreciente de conjuntos cerrados (no vacíos) cuya intersección sea vacía.

Ejercicio 22. Sea X un espacio vectorial normado, sean $A, B \subset X$ cerrados y sean $C, D \subset X$ compactos. Probar que

1. Si $d(C,D) = \inf_{\boldsymbol{x}\in C,\boldsymbol{y}\in D} \|\boldsymbol{x}-\boldsymbol{y}\| = 0$ entonces $C \cap D \neq \varnothing$.
2. Si $d(A,B) = \inf_{\boldsymbol{x}\in C,\boldsymbol{y}\in D} \|\boldsymbol{x}-\boldsymbol{y}\| = 0$ entonces no necesariamente $A \cap B \neq \varnothing$.

Ejercicio 23. Sean $(X, \|\cdot\|_X)$ y $(Y, \|\cdot\|_Y$ dos espacios vectoriales normados. Definimos $(X\times Y, \|\cdot\|)$ como el espacio vectorial normado donde $(\boldsymbol{x},\boldsymbol{y}) \in X\times Y \iff \boldsymbol{x}\in X \wedge \boldsymbol{y}\in Y$ y la norma en $X\times Y$ se define como $\|(\boldsymbol{x},\boldsymbol{y})\| = \|\boldsymbol{x}\|_X + \|\boldsymbol{y}\|_Y$. Sean $A\subset X$ y $B \subset Y$ compactos. Demuestre que entonces $A\times B$ es compacto en $X\times Y$.

Ejercicio 24. Sea $U = \mathbb{R}^{\mathbb{N}}$, es decir, el espacio de las sucesiones. Considere en U la norma

$$\|\boldsymbol{x}\| = \sup\{|x_i| : i \in \mathbb{N}\}$$

Indique cual(es) de los siguientes conjuntos son compactos:

1. $B_0(0,1) = \{\boldsymbol{x}\in U : |x_i| \le 1, \forall i \in \mathbb{N}\}$
2. $B_1(0,1) = \{\boldsymbol{x}\in U : |x_i| \le \frac{1}{i}, \forall i \in \mathbb{N}\}$
3. $B_2(0,1) = \{\boldsymbol{x}\in U : |x_i| = 1, \forall i \in \mathbb{N}\}$

Ejercicio 25. Demuestre que las matrices ortogonales de $n\times n$ forman un subconjunto compacto de $\mathbb{R}^n\times\mathbb{R}^n$.

CAPÍTULO 5

Teoremas de la Función Inversa e Implícita

Estudiaremos algunos teoremas que nos permiten resolver sistemas de ecuaciones no lineales, conocer las condiciones para que exista la inversa de una función de varias variables y bajo qué condiciones podemos expresar una o varias variables de una función en términos de las demás (o en términos de una variable conocida).

5.1. Teorema del punto fijo de Banach

Motivación: Los teoremas de punto fijo nos sirven para demostrar la existencia de soluciones de una ecuación, la existencia de equilibrio en un sistema, encontrar raíces de funciones o resolver sistemas de ecuaciones no lineales.

Problema 5.1. Encontrar una solución a:

$$\begin{cases} u' &= f(x,u) \\ u(x_0) &= u_0 \end{cases}$$

donde $x, u \in \mathbb{R}$. Este problema es equivalente a

$$u(x) = u_0 + \int_{x_0}^{x} f(s, u(s))ds$$

Supongamos que $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ es continua, entonces

$$\begin{aligned} T : C([x_0 - a, x_0 + a], \mathbb{R}) &\to C([x_0 - a, x_0 + a], \mathbb{R}) \\ u &\mapsto u_0 + \int_{x_0}^{x} f(s, u(s))ds \end{aligned}$$

está bien definida y el problema original es equivalente a encontrar u tal que

$$u = T(u)$$

que recibe el nombre de problema de punto fijo.

Problema 5.2. Dado $F:\mathbb{R}^n\to\mathbb{R}^n$ encontrar solución a:

$$F(\boldsymbol{x})=0$$

Este es un problema de punto fijo, pues se puede escribir

$$\boldsymbol{x}=\boldsymbol{x}+F(\boldsymbol{x})$$

$$\begin{array}{rccc} T: & \mathbb{R}^n & \to & \mathbb{R}^n \\ & \boldsymbol{x} & \mapsto & \boldsymbol{x}+F(\boldsymbol{x}) \end{array}$$

Definiendo $T(\boldsymbol{x})=\boldsymbol{x}+F(\boldsymbol{x})$ el problema original consiste en encontrar un punto fijo de T, es decir en encontrar un $\boldsymbol{x}$ tal que $T(\boldsymbol{x})=\boldsymbol{x}$.

En diversas áreas de la ingeniería es común encontrarse con problemas donde se quiere resolver

$$u=T(u),\ \text{con } T:E\to E$$

o bien $T:K\to K$ en que $K\subseteq E$ es un subconjunto cerrado y E es un espacio de Banach.

Definición 5.1. $T:K\to K$ se dice contracción si existe una constante c , $0<c<1$ tal que

$$\|T(\boldsymbol{u})-T(\boldsymbol{v})\|\leq c\|\boldsymbol{u}-\boldsymbol{v}\|\ \forall \boldsymbol{u},\boldsymbol{v}\in K$$

Teorema 5.1. (Teorema del punto fijo de Banach)
Sea E espacio de Banach, $K\subseteq E$ cerrado. Si $T:K\to K$ es una contracción entonces existe un y solo un punto fijo de T en K

$$\exists!\boldsymbol{x}\in K,\ T(\boldsymbol{x})=\boldsymbol{x}$$

Demostración. Veamos primero la existencia. El método de demostración es constructivo.
Sea $\boldsymbol{x}_0\in K$ cualquiera. Consideremos la sucesión $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ definida por la recurrencia

$$\boldsymbol{x}_{k+1}=T(\boldsymbol{x}_k),\ k\in\mathbb{N}$$

Demostremos que $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ es de Cauchy.

$$\begin{aligned} \|\boldsymbol{x}_k-\boldsymbol{x}_m\| &= \|T(\boldsymbol{x}_{k-1})-T(\boldsymbol{x}_{m-1})\| \\ &\leq c\|\boldsymbol{x}_{k-1}-\boldsymbol{x}_{m-1}\| \\ &\leq c\|T(\boldsymbol{x}_{k-2})-T(\boldsymbol{x}_{m-2})\| \\ &\leq c^2\|\boldsymbol{x}_{k-2}-\boldsymbol{x}_{m-2}\| \end{aligned}$$

supongamos, sin perdida de generalidad que $k\geq m$. Si repetimos el procedimiento anterior, obtenemos

$$\|\boldsymbol{x}_k-\boldsymbol{x}_m\|\leq c^m\|\boldsymbol{x}_{k-m}-\boldsymbol{x}_0\|$$

en particular

$$\|\boldsymbol{x}_{l+1}-\boldsymbol{x}_m\|\leq c^m\|\boldsymbol{x}_1-\boldsymbol{x}_0\|$$

Entonces podemos escribir

$$\begin{aligned} \|\boldsymbol{x}_k-\boldsymbol{x}_m\| &\leq \|\boldsymbol{x}_k-\boldsymbol{x}_{k-1}\|+\|\boldsymbol{x}_{k-1}-\boldsymbol{x}_{k-2}\|+\cdots+\|\boldsymbol{x}_{m+1}-\boldsymbol{x}_m\| \\ &\leq (c^{k-1}+c^{k-2}+\cdots+c^m)\|\boldsymbol{x}_1-\boldsymbol{x}_0\| \end{aligned}$$

es decir,

$$\begin{aligned} \|\boldsymbol{x}_k - \boldsymbol{x}_m\| &\leq \sum_{i=m}^{\infty} c^i \|\boldsymbol{x}_1 - \boldsymbol{x}_0\| \\ &= c^m \left(\sum_{i=0}^{\infty} c^i \right) \|\boldsymbol{x}_1 - \boldsymbol{x}_0\| \\ &= c^m \frac{1}{1-c} \|\boldsymbol{x}_1 - \boldsymbol{x}_0\| \overset{k,m\to\infty}{\to} 0 \end{aligned}$$

Por lo tanto $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ es una sucesión de Cauchy en E, luego tiene un límite $\boldsymbol{x}^* \in E$ y como K es cerrado entonces $\boldsymbol{x}^* \in K$. Por otra parte, si T es contracción, es continua (*tarea*), por lo que tomando límite en la ecuación $\boldsymbol{x}_{k+1} = T(\boldsymbol{x}_k)$ obtenemos que

$$\boldsymbol{x}^* = T(\boldsymbol{x}^*)$$

es decir, $\boldsymbol{x}^*$ es un punto fijo de T.
Veamos que es único. Supongamos que T tiene dos puntos fijos, $\boldsymbol{x} = T(\boldsymbol{x})$ e $\boldsymbol{y} = T(\boldsymbol{y})$, entonces

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \|T(\boldsymbol{x}) - T(\boldsymbol{y})\| \leq c\|\boldsymbol{x} - \boldsymbol{y}\|$$

y como $0 \leq c < 1$, no queda otra posibilidad más que $\boldsymbol{x} = \boldsymbol{y}$. ■

Nota 5.1. El teorema del punto fijo de Banach provee una condición suficiente (y no necesaria) para la existencia de puntos fijos. No es necesario que una función sea contractante para que tenga punto fijo, un claro ejemplo es la función $f : [0,1] \to [0,1]$ definida por $f(x) = x$ la cual es continua, no contractante y cumple que cualquier $x \in [0,1]$ es un punto fijo de la función.

Ejemplo 5.1. (Existencia de soluciones para EDO)
Volvamos al problema 5.1. Si $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ es continua y satisface

$$|f(s,u) - f(s,v)| \leq K|u - v|$$

entonces si $x > x_0$

$$\begin{aligned} |T(u) - T(v)| &= \left| \int_{x_0}^{x} f(s,u(s)) - f(s,v(s)) ds \right| \\ &\leq \int_{x_0}^{x} |f(s,u(s)) - f(s,v(s))| ds \\ &\leq K \int_{x_0}^{x} |u(s) - v(s)| ds \\ &\leq K\|u - v\|_\infty a \end{aligned}$$

si $x < x_0$ se obtiene el mismo resultado, luego $\|T(u) - T(v)\|_\infty \leq Ka\|u - v\|_\infty$ para toda $u, v \in C([x_0 - a, x_0 + a])$. Si $Ka < 1$ entonces, gracias al teorema del punto fijo de Banach, existe $u \in C([x_0 - a, x_0 + a])$ tal que

$$u = T(u)$$

y este u es único.
Para encontrar una solución se puede iterar, reproduciendo la demostración del teorema de punto fijo de Banach.

$$\begin{aligned} u_{n+1} &= Tu_n \\ u_{n+1}(x) &= u_0 + \int_{x_0}^{x} f(s, u_n(s)) ds \end{aligned}$$

este método para encontrar la solución recibe el nombre de método de Picard.

Ejemplo 5.2. (Existencia de soluciones de una ecuación)
Consideremos el sistema de ecuaciones

$$\begin{aligned} x &= \frac{1}{2}\cos(x+y) \\ y &= \frac{1}{3}\ln(1+x^2+y^2)+5 \end{aligned}$$

este es un problema de punto fijo: $(x,y)=T(x,y)$ con $T:\mathbb{R}^2\to\mathbb{R}$ definida por

$$T(x,y)=\left(\frac{1}{2}\cos(x+y),\frac{1}{3}\ln(1+x^2+y^2)+5\right)$$

Vamos a demostrar que T es contractante. Como consecuencia, gracias al teorema del punto fijo de Banach, tendremos que el sistema de ecuaciones tiene una y solo una solución en $\mathbb{R}^2$. Para ello recordemos el teorema del valor medio: Sea $f:\mathbb{R}^n\to\mathbb{R}$ diferenciable, entonces

$$f(x)-f(y)=\int_0^1 \nabla f(tx+(1-t)y)\cdot(x-y)dt$$

lo que implica que

$$\begin{aligned} |f(x)-f(y)| &\leq \int_0^1 \|\nabla f(tx+(1-t))y\|\|x-y\|dt \\ &\leq \max_{0\leq t\leq 1}\{\|\nabla f(tx+(1-t)y)\|\}\|x-y\| \end{aligned}$$

Supongamos que $\|\nabla f(z)\|\leq K$ para todo $z\in\mathbb{R}^n$. Entonces $|f(x)-f(y)|\leq K\|x-y\|$.
Apliquemos esto a nuestro problema. Si $T_1(x,y)=(1/2)\cos(x+y)$ entonces

$$\|\nabla T_1(x,y)\|=\left\|\left(-\frac{1}{2}\operatorname{sen}(x+y),-\frac{1}{2}\operatorname{sen}(x+y)\right)\right\|\leq\frac{1}{\sqrt{2}}$$

$$\Rightarrow\ |T_1(x,y)-T_1(\bar{x},\bar{y})|\leq\frac{1}{\sqrt{2}}\cdot\|(x,y)-(\bar{x},\bar{y})\|$$

Para $T_2(x,y)=(1/3)\ln(1+x^2+y^2)+5$ hacemos lo mismo:

$$\|\nabla T_2(x,y)\|=\frac{1}{3}\left\|\left(\frac{2x}{1+x^2+y^2},\frac{2y}{1+x^2+y^2}\right)\right\|\leq\frac{2}{3}\sqrt{\left(\frac{x}{1+x^2}\right)^2+\left(\frac{y}{1+y^2}\right)^2}$$

la función $f(z)=z/(1+z^2)$ alcanza su máximo en $z=1$ (*tarea*) y su máximo es $1/2$. Por lo tanto $\|\nabla T_2(x,y)\|\leq\sqrt{2}/3$ lo que implica

$$|T_2(x,y)-T_2(\bar{x},\bar{y})|\leq\frac{\sqrt{2}}{3}\cdot\|(x,y)-(\bar{x},\bar{y})\|$$

Concluimos entonces que

$$\begin{aligned} \|T(x,y)-T(\bar{x},\bar{y})\| &= \sqrt{(T_1(x,y)-T_1(\bar{x},\bar{y}))^2+(T_2(x,y)-T_2(\bar{x},\bar{y}))^2} \\ &\leq \sqrt{\frac{13}{18}}\cdot\|(x,y)-(\bar{x},\bar{y})\| \end{aligned}$$

es decir, T es contractante pues $\sqrt{13/18}<1$.

5.2. Teorema de la función inversa

Motivación: Sea $L : \mathbb{R}^n \to \mathbb{R}^n$ una función lineal. Para que L sea biyectiva es necesario y suficiente que la matriz asociada (matriz representante) sea invertible, y en este caso la matriz representante de la función inversa será la inversa de la matriz representante de L.
Si $L(\boldsymbol{x}_0) = \boldsymbol{b}_0$ y queremos resolver la ecuación $L(\boldsymbol{x}) = \boldsymbol{b}_0 + \Delta\boldsymbol{b}$, la solución será

$$\boldsymbol{x} = \boldsymbol{x}_0 + \Delta\boldsymbol{x} = L^{-1}(\boldsymbol{b}_0) + L^{-1}(\Delta\boldsymbol{b})$$

Cuando $F : \Omega \to \mathbb{R}^n$ es diferenciable, F es localmente "como" una función lineal (afín), y por lo tanto es razonable pensar que la biyectividad local de F esté relacionada con la biyectividad de su aproximación lineal.

Teorema 5.2. (Teorema de la función inversa)
Sea $F : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una función de clase $\mathcal{C}^1$. Supongamos que para $\boldsymbol{a} \in \Omega$, $DF(\boldsymbol{a})$ (Jacobiano) es invertible y que $F(\boldsymbol{a}) = \boldsymbol{b}$. Entonces

1. *Existen abiertos U y V de $\mathbb{R}^n$ tales que $\boldsymbol{a} \in U$, $\boldsymbol{b} \in V$ y $F : U \to V$ es biyectiva.*
2. *Si $G : V \to U$ es la inversa de F, es decir, $G = F^{-1}$, entonces G es también de clase $\mathcal{C}^1$ y $DG(\boldsymbol{b}) = [DF(\boldsymbol{a})]^{-1}$.*

Antes de dar la demostración veremos algunos ejemplos.

Ejemplo 5.3. Sea $F(x, y) = (x^2 + \ln y, y^2 + xy^3)$, entonces $F(1, 1) = (1, 2)$
Consideremos la ecuación $F(x, y) = (b_1, b_2)$

$$\begin{aligned} x^2 + \ln y &= b_1 \\ y^2 + xy^3 &= b_2 \end{aligned}$$

con (b_1, b_2) "cerca" de $(1, 2)$.

$$DF(x, y) = \begin{pmatrix} 2x & 1/y \\ y^3 & 2y + 3xy^2 \end{pmatrix}$$

evaluando el Jacobiano en $(x, y) = (1, 1)$ obtenemos

$$DF(1, 1) = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$$

que es una matriz invertible.
Concluimos entonces, gracias al teorema de la función inversa que la ecuación tiene una y solo una solución

$$(x(b_1, b_2), y(b_1, b_2))$$

para cualquier (b_1, b_2) cercano al punto $(1, 2)$. Más aún $(x(b_1, b_2), y(b_1, b_2))$ está cerca de $(1, 1)$ y depende de (b_1, b_2) de manera diferenciable.

Ejemplo 5.4. Cambio de coordenadas esféricas:

$$\begin{aligned} x &= r \operatorname{sen}\phi \cos\theta \\ y &= r \operatorname{sen}\phi \operatorname{sen}\theta \\ z &= r \cos\phi \end{aligned}$$

Llamaremos Φ al cambio de variables. El Jacobiano de la transformación en el punto $r = 1, \phi = \pi/4, \theta = \pi/4$ es

$$D\Phi(1, \pi/4, \pi/4) = \begin{pmatrix} 1/2- & 1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 \\ \sqrt{2}/2 & 0- & \sqrt{2}/2 \end{pmatrix}$$

Esta matriz es invertible pues su determinante es igual a $\sqrt{2}/2$. En general se tiene que $|D\Phi(r, \theta, \phi)| = r^2 \operatorname{sen} \phi$ que es distinto de cero excepto si $r = 0$ o si $\operatorname{sen} \phi = 0$. Por lo tanto, salvo en el eje z la transformación Φ es localmente biyectiva.

Previo a la demostración del teorema, consideremos tres resultados que serán de utilidad.

1. A partir de la norma descrita en (1.10), consideremos en $M_{n\times n}(\mathbb{R})$ la norma

$$\|A\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2}$$

Entonces, $\|A\boldsymbol{x}\|_2 \leq \|A\|_F \|\boldsymbol{x}\|_2$ si $\boldsymbol{x} \in \mathbb{R}^n$, $A \in M_{n\times n}$ y además $\|AB\|_F \leq \|A\|_F \|B\|_F$, si $A, B \in M_{n\times n}$.

2. A partir del teorema del punto fijo de Banach (teorema 5.1) consideremos el siguiente corolario: Si $(E, \|\cdot\|)$ es un espacio de Banach,$C \subseteq E$ es un conjunto cerrado y $F : C \to C$ es una función contractante, entonces existe un único $\boldsymbol{x} \in C$ tal que $F(\boldsymbol{x}) = \boldsymbol{x}$.

3. A partir del teorema del valor medio (teorema 1.23) y de la caracterización de convexidad (definición 4.16): Sean $F : U \to \mathbb{R}^n$ de clase $\mathcal{C}^1$ y U un conjunto convexo. Si $\|DF_i(\boldsymbol{x})\| \leq A_i$ $\forall \boldsymbol{x} \in U$, entonces $\|F(\boldsymbol{x}) - F(\boldsymbol{y})\| \leq \sqrt{A_1^2 + ... + A_n^2} \|\boldsymbol{x} - \boldsymbol{y}\|$ $\forall \boldsymbol{x}, \boldsymbol{y} \in U$.
Si $\|DF(\boldsymbol{x})\| \leq C$ $\forall \boldsymbol{x} \in U$ entonces $\|F(\boldsymbol{x}) - F(\boldsymbol{y})\| \leq C\|\boldsymbol{x} - \boldsymbol{y}\|$, pues

$$\begin{aligned} \|F(\boldsymbol{x}) - F(\boldsymbol{y})\| &= \left\| \int_0^1 DF(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) dt \right\| \\ &\leq \int_0^1 \|DF(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y})\| dt \\ &\leq \int_0^1 \|DF(t\boldsymbol{x} + (1-t)\boldsymbol{y})\| \|\boldsymbol{x} - \boldsymbol{y}\| dt \\ &\leq C\|\boldsymbol{x} - \boldsymbol{y}\| \end{aligned}$$

Demostración. (Teorema de la función inversa)
Observemos que si F es de clase $\mathcal{C}^1$ entonces la función

$$\begin{aligned} DF : \Omega &\to M_{n\times n} \\ \boldsymbol{x} &\mapsto DF(\boldsymbol{x}) \end{aligned}$$

es continua pues

$$\|DF(\boldsymbol{x}) - DF(\boldsymbol{x}_0)\|_F = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \left| \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}) - \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0) \right|^2}$$

y como todas las derivadas parciales son continuas, dado $\varepsilon > 0$, existen $\delta_{ij} > 0$ tales que $\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta_{ij} \Rightarrow |\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}) - \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)| < \frac{\varepsilon}{n}$, entonces escogiendo $\delta = \min_{i,j} \delta_{ij} > 0$ tenemos que

$\|\boldsymbol{x}-\boldsymbol{x}_0\| < \delta \Rightarrow \|DF(\boldsymbol{x})-DF(\boldsymbol{x}_0)\|_F < \sqrt{n^2(\frac{\varepsilon}{n})^2} = \varepsilon$.
Queremos demostrar que existen abiertos U y V tales que $F: U \to V$ es biyectiva, lo que se puede expresar en otras palabras como: Dado $\boldsymbol{y} \in V$ existe un único $\boldsymbol{x} \in U$ que es solución de la ecuación $F(\boldsymbol{x}) = \boldsymbol{y}$. Además

$$\begin{array}{rcll} F(\boldsymbol{x}) = \boldsymbol{y} & \Leftrightarrow & \boldsymbol{y} - F(\boldsymbol{x}) = 0 & \\ & \Leftrightarrow & DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - F(\boldsymbol{x})) = 0 & \text{pues } DF(\boldsymbol{a}) \text{ es invertible} \\ & \Leftrightarrow & \boldsymbol{x} + DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - F(\boldsymbol{x})) = \boldsymbol{x} & \end{array}$$

Entonces dado $\boldsymbol{y} \in V$, resolver la ecuación $F(\boldsymbol{x}) = \boldsymbol{y}$ es equivalente a encontrar un punto fijo a la función $\varphi_y(\boldsymbol{x}) = \boldsymbol{x} + DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - F(\boldsymbol{x}))$ Como DF es continua en $\boldsymbol{a}$, existe $\varepsilon > 0$ tal que

$$\boldsymbol{x} \in B(\boldsymbol{a}, \varepsilon) \Rightarrow \|DF(\boldsymbol{x}) - DF(\boldsymbol{a})\|_F < \frac{1}{2\sqrt{n}\|DF(\boldsymbol{a})^{-1}\|_F}$$

Para probar que $F : B(\boldsymbol{a}, \varepsilon) \to \mathbb{R}^n$ es inyectiva, basta probar que φ_y es contractante, en efecto $F(\boldsymbol{x}_1) = \boldsymbol{y} \wedge F(\boldsymbol{x}_2) = \boldsymbol{y} \Rightarrow \varphi_y(\boldsymbol{x}_1) = \boldsymbol{x}_1 \wedge \varphi_y(\boldsymbol{x}_2) = \boldsymbol{x}_2$ y si φ_y es contractante entonces $\|\varphi_y(\boldsymbol{x}_1) - \varphi_y(\boldsymbol{x}_2)\| \leq C\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \Rightarrow \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leq C\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|$ y como $C < 1$, $\boldsymbol{x}_1 = \boldsymbol{x}_2$.
Observemos que φ_y es de clase $\mathcal{C}^1$. Para probar que es contractante acotamos la norma de su derivada:

$$D\varphi_y(\boldsymbol{x}) = I - DF(\boldsymbol{a})^{-1}DF(\boldsymbol{x}) = DF(\boldsymbol{a})^{-1}(DF(\boldsymbol{a}) - DF(\boldsymbol{x}))$$

donde I es la matriz identidad de $n \times n$.

$$\|D\varphi_y(\boldsymbol{x})\|_F \leq \|DF(\boldsymbol{a})^{-1}\|_F\|DF(\boldsymbol{a}) - DF(\boldsymbol{x})\|_F \leq \frac{1}{2\sqrt{n}}$$

la última desigualdad es válida si $\boldsymbol{x} \in B(\boldsymbol{a}, \varepsilon)$.
Entonces como $B(\boldsymbol{a}, \varepsilon)$ es convexa y $\|D(\varphi_y)_i(\boldsymbol{x})\| \leq \frac{1}{2\sqrt{n}}$ para $i = 1, ..., n$, obtenemos

$$\|\varphi_y(\boldsymbol{x}) - \varphi_y(\boldsymbol{z})\| \leq \|\boldsymbol{x} - \boldsymbol{z}\|\sqrt{\sum_{i=1}^{n}(\frac{1}{2\sqrt{n}})^2} = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{z}\| \ \forall \boldsymbol{x}, \boldsymbol{z} \in B(\boldsymbol{a}, \varepsilon)$$

Si definimos $U = B(\boldsymbol{a}, \varepsilon)$ y $V = F(U)$, entonces $F : U \to V$ es biyectiva. Veamos que V es abierto. Sea $\boldsymbol{y}^* \in V$ y $\boldsymbol{x}^* \in B(\boldsymbol{a}, \varepsilon)$ tal que $F(\boldsymbol{x}^*) = \boldsymbol{y}^*$. Hay que demostrar que existe $\rho > 0$ tal que si $\boldsymbol{y} \in B(\boldsymbol{y}^*, \rho)$ entonces existe $\boldsymbol{x} \in U$ tal que $F(\boldsymbol{x}) = \boldsymbol{y} \Leftrightarrow \boldsymbol{x} = \varphi_y(\boldsymbol{x})$, es decir, $B(\boldsymbol{y}^*, \rho) \subseteq F(U) = V$.
Sea $r > 0$ tal que $B(\boldsymbol{x}^*, 2r) \subseteq U$, y $\boldsymbol{x} \in \bar{B}(\boldsymbol{x}^*, r)$

$$\begin{array}{rcl} \varphi_y(\boldsymbol{x}) - \boldsymbol{x}^* & = & \varphi_y(\boldsymbol{x}) - (\boldsymbol{x}^* + DF(\boldsymbol{a})^{-1})\boldsymbol{y} - F(\boldsymbol{x}^*))) + DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - \boldsymbol{y}^*) \\ & = & \varphi_y(\boldsymbol{x}) - \varphi_y(\boldsymbol{x}^*) + DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - \boldsymbol{y}^*) \end{array}$$

y por lo tanto

$$\begin{array}{rcl} \|\varphi_y(\boldsymbol{x}) - \boldsymbol{x}^*\| & \leq & \|\varphi_y(\boldsymbol{x}) - \varphi_y(\boldsymbol{x}^*)\| + \|DF(\boldsymbol{a})^{-1}(\boldsymbol{y} - \boldsymbol{y}^*\|) \\ & \leq & \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^*\| + \|DF(\boldsymbol{a})^{-1}\|\|\boldsymbol{y} - \boldsymbol{y}^*\| \end{array}$$

Si escogemos $\rho = r/(2\|DF(\boldsymbol{a})^{-1}\|_F)$ tendremos que $\forall \boldsymbol{y} \in B(\boldsymbol{y}^*, \rho)$ la función

$$\varphi_y : \bar{B}(\boldsymbol{x}^*, r) \to \bar{B}(\boldsymbol{x}^*, r)$$

es una contracción y por lo tanto tiene un único punto fijo $\boldsymbol{x} \in \bar{B}(\boldsymbol{x}^*, r) \subseteq U$ tal que

$$\varphi_y(\boldsymbol{x}) = \boldsymbol{x} \Rightarrow F(\boldsymbol{x}) = \boldsymbol{y}$$

Concluimos así que $B(\boldsymbol{y}^*, \rho) \subseteq V$, es decir, V es abierto.
Resta estudiar la diferenciabilidad de la función inversa. Sean $\boldsymbol{y}, \boldsymbol{y}+\boldsymbol{k} \in V$ entonces existen únicos $\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{h} \in U$ tales que $\boldsymbol{y} = F(\boldsymbol{x})$ e $\boldsymbol{y}+\boldsymbol{k} = F(\boldsymbol{x}+\boldsymbol{h})$ entonces

$$\begin{aligned} F^{-1}(\boldsymbol{y}+\boldsymbol{k}) - F^{-1}(\boldsymbol{y}) - DF(\boldsymbol{x})^{-1}\boldsymbol{k} &= \boldsymbol{h} - DF(\boldsymbol{x})^{-1}\boldsymbol{k} \\ &= DF(\boldsymbol{x})^{-1}(F(\boldsymbol{x}+\boldsymbol{h}) - F(\boldsymbol{x}) - DF(\boldsymbol{x})\boldsymbol{h}) \end{aligned}$$

Por lo tanto

$$\frac{1}{\|\boldsymbol{k}\|}\|F^{-1}(\boldsymbol{y}+\boldsymbol{k}) - F^{-1}(\boldsymbol{y}) - DF(\boldsymbol{x})^{-1}\boldsymbol{k}\| \leq \|DF(\boldsymbol{x})^{-1}\| \frac{\|F(\boldsymbol{x}+\boldsymbol{h}) - F(\boldsymbol{x}) - DF(\boldsymbol{x})\boldsymbol{h}\|}{\|\boldsymbol{h}\|} \frac{\|\boldsymbol{h}\|}{\|\boldsymbol{k}\|}$$

Probaremos ahora que $\|\boldsymbol{h}\|/\|\boldsymbol{k}\| \leq C < \infty$ lo que implica que $\boldsymbol{h} \to 0$ cuando $\boldsymbol{k} \to 0$ y como F es diferenciable en el punto $\boldsymbol{x}$, el lado derecho de la desigualdad anterior tiende a cero cuando $\boldsymbol{k} \to 0$, obligando asi a que el lado izquierdo tienda a cero también, lo que por definición significa que F^{-1} es diferenciable en $\boldsymbol{y}$ y

$$DF^{-1}(\boldsymbol{y}) = [DF(\boldsymbol{x})]^{-1}$$

Se sabe que

$$\begin{aligned} \|\boldsymbol{h} - DF(\boldsymbol{a})^{-1}\boldsymbol{k}\| &= \|\boldsymbol{h} - DF(\boldsymbol{a})^{-1}(F(\boldsymbol{x}+\boldsymbol{h}) - F(\boldsymbol{x}))\| \\ &= \|\boldsymbol{h} + \boldsymbol{x} - DF(\boldsymbol{a})^{-1}(F(\boldsymbol{x}+\boldsymbol{h}) - \boldsymbol{y}) - \boldsymbol{x} + \\ &\quad DF(\boldsymbol{a})^{-1}(F(\boldsymbol{x}) - \boldsymbol{y})\| \\ &= \|\varphi_y(\boldsymbol{x}+\boldsymbol{h}) - \varphi_y(\boldsymbol{x})\| \\ &\leq \frac{1}{2}\|\boldsymbol{h}\| \end{aligned}$$

por lo tanto

$$\|\boldsymbol{h}\| - \|DF(\boldsymbol{a})^{-1}\| \leq \|\boldsymbol{h} - DF(\boldsymbol{a})^{-1}\| \leq 1/2\|\boldsymbol{h}\|$$

lo que implica que

$$1/2\|\boldsymbol{h}\| \leq \|DF(\boldsymbol{a})^{-1}\boldsymbol{k}\| \leq \|DF(\boldsymbol{a})^{-1}\|\|\boldsymbol{k}\|$$

es decir,

$$\frac{\|\boldsymbol{h}\|}{\|\boldsymbol{k}\|} \leq 2\|DF(\boldsymbol{a})^{-1}\| < \infty$$

que es lo que se quería probar.
La continuidad de $DF^{-1}(\boldsymbol{y}) = [DF(F^{-1}(\boldsymbol{y}))]^{-1}$ se obtiene de la regla de Cramer para obtener la inversa de una matriz. Para una matriz invertible A se tiene que

$$A^{-1} = \frac{1}{\det(A)}[\text{cofactores}(A)]^T$$

y donde "cofactores(A)" es una matriz formada por los distintos subdeterminantes de A que se obtienen al sacarle una fila y una columna a A. Entonces como F es continuamente diferenciable, sus derivadas parciales son continuas, y gracias a la formula de Cramer las derivadas parciales de F^{-1} serán continuas también pues son productos y sumas de las derivadas parciales de F y cuociente con el determinante de $DF(F^{-1}(\boldsymbol{y}))$ que es distinto de cero y continuo por la misma razón. ■

5.3. Teorema de la función implícita

Motivación: Supongamos que se tiene una función $f : \mathbb{R}^2 \to \mathbb{R}$, y consideremos una ecuación de la forma

$$f(x, y) = 0$$

Es entonces natural hacer la siguiente pregunta: ¿Es posible despejar y en función de x?
Supongamos que sí, es decir, existe una función $y(x)$ que satisface

$$f(x, y(x)) = 0$$

supongamos además que f e $y(x)$ son diferenciables, entonces podemos derivar la ecuación anterior con respecto a x:

$$\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0$$

o sea

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} \quad \text{(Derivación implícita)}$$

Notemos que para poder realizar este cálculo es necesario que $\partial f/\partial y$ sea distinto de cero.

Ejemplo 5.5. Consideremos la siguiente ecuación

$$x^2 + y^2 = 1$$

¿Es posible despejar y en función de x? Evidentemente la respuesta a esta pregunta es que no es posible despejar globalmente una variable en función de la otra, pero si (x_0, y_0) es un punto que satisface la ecuación, es decir, $x_0^2 + y_0^2 = 1$ y además $y_0 \neq 0$, entonces es posible despejar localmente y en función de x (es decir, en una vecindad de (x_0, y_0)). Además derivando la ecuación se tiene que

$$2x + 2y\frac{dy}{dx} = 0 \;\Rightarrow\; \frac{dy}{dx} = -\frac{x}{y}$$

Sin embargo si $y_0 = 0$ es imposible despejar y en función de x en torno al punto (x_0, y_0).

Ejemplo 5.6. Consideremos ahora un caso más general que el anterior. Suponga que se tienen las variables $x_1, ..., x_m$ e $y_1, ..., y_n$ relacionadas por n ecuaciones (sistema de ecuaciones),
$F_i(x_1, ...x_m, y_1, ...y_n) = 0,\ i : 1, ..., n$.Esto se puede escribir como

$$\begin{aligned} F : \mathbb{R}^m \times \mathbb{R}^n &\to \mathbb{R}^n \\ (x_1, .., x_m, y_1, ..., y_n) &\mapsto F(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} F_1(x_1, ..., x_m, y_1, ..., y_n) \\ \vdots \\ F_n(x_1, ..., x_m, y_1, ..., y_n) \end{pmatrix} \end{aligned}$$

Entonces, $F(x_1, ..., x_m, y_1, ..., y_n) = 0$.
¿Es posible despejar las variables $\boldsymbol{y}$, en función de las variables $\boldsymbol{x}$? Es decir, existen funciones $y_j(x_1, ..., x_m),\ j = 1, ..., n$ tales que

$$F(x_1, ..., x_m, y_1(x_1, ..., x_x), ..., y_n(x_1, ..., x_m)) = 0$$

Veamos el caso particular de un sistema lineal. Sea L una matriz $L = [A\ B] \in M_{n\times(n+m)}$, y consideremos el sistema

$$A\boldsymbol{x} + B\boldsymbol{y} = \boldsymbol{b}$$

En este caso es posible despejar $\boldsymbol{y}$ en función de $\boldsymbol{x}$ cuando B es invertible:

$$\boldsymbol{y}(\boldsymbol{x}) = B^{-1}\boldsymbol{b} - B^{-1}A\boldsymbol{x}$$

Teorema 5.3. (Teorema de la función implícita)
Sea $F : \Omega \subseteq \mathbb{R}^{m+n} \to \mathbb{R}^n$ *una función de clase* $\mathcal{C}^1$*, y sea* $(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{R}^m \times \mathbb{R}^n$ *un punto tal que* $F(\boldsymbol{a}, \boldsymbol{b}) = 0$*. Escribimos entonces* $DF(\boldsymbol{a}, \boldsymbol{b}) = [D_x F(\boldsymbol{a}, \boldsymbol{b})\, D_y F(\boldsymbol{a}, \boldsymbol{b})]$ *y supongamos que* $D_y F(\boldsymbol{a}, \boldsymbol{b})$ *es invertible. Entonces existen conjuntos abiertos* $U \subseteq \mathbb{R}^{m+n}, W \subseteq \mathbb{R}^m$ *con* $(\boldsymbol{a}, \boldsymbol{b}) \in U$ *y* $\boldsymbol{a} \in W$ *tales que para cada* $\boldsymbol{x} \in W$ *existe un único* $\boldsymbol{y}$ *tal que* $(\boldsymbol{x}, \boldsymbol{y}) \in U$ *y*

$$F(\boldsymbol{x}, \boldsymbol{y}) = 0$$

esto define una función $G : W \to \mathbb{R}^n$ *que es de clase* $\mathcal{C}^1$ *y que satisface*

$$F(\boldsymbol{x}, G(\boldsymbol{x})) = 0, \ \forall \boldsymbol{x} \in W$$

además $DG(\boldsymbol{x}) = -[D_y F(\boldsymbol{x}, G(\boldsymbol{x}))]^{-1} D_x F(\boldsymbol{x}, G(\boldsymbol{x}))$ *para todo* $\boldsymbol{x} \in W$ *y* $G(\boldsymbol{a}) = \boldsymbol{b}$*.*

Al igual que en el teorema de la función inversa veamos un ejemplo antes de dar la demostración del teorema.

Ejemplo 5.7. Considere el sistema de ecuaciones

$$x^2 + \operatorname{sen}(y) + \cos(yz) - w^3 - 1 = 0$$

$$x^3 + \cos(yx) + \operatorname{sen}(z) - w^2 - 1 = 0$$

1. Muestre que es posible despejar (y, z) en función de (x, w) en una vecindad del punto

   $$(x_0, y_0, z_0, w_0) = (0, 0, 0, 0)$$

   Calcular $\frac{\partial y}{\partial x}(0, 0)$.

2. ¿Es posible despejar (x, y) en función de las variables (z, w) en una vecindad de $(0, \pi, 0, 0)$? ¿Qué se puede decir de despejar (x, w) en función de (y, z)?

Solución.

1. Para este problema se tiene que la función F es:
   $F(x, y, z, w) = (x^2 + \operatorname{sen}(y) + \cos(yz) - w^3 - 1, x^3 + \cos(yx) + \operatorname{sen}(z) - w^2 - 1)$.
   Entonces F es de clase $\mathcal{C}^1$ y $F(0, 0, 0, 0) = (0, 0)$. Además

   $$D_{(y,z)} F(x, y, z, w) = \begin{pmatrix} \partial F_1/\partial y & \partial F_1/\partial z \\ \partial F_2/\partial y & \partial F_2/\partial z \end{pmatrix} = \begin{pmatrix} \cos y - \operatorname{sen}(yz)z & -\operatorname{sen}(yz)y \\ -\operatorname{sen}(yx)x & \cos(z) \end{pmatrix}$$

   luego

   $$D_{(y,z)} F(0, 0, 0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

   que es invertible. Luego, gracias al teorema, es posible despejar (y, z) en función de (x, w) de manera diferenciable en una vecindad del punto $(0, 0)$. Si derivamos las ecuaciones con respecto a x obtenemos

   $$2x + \cos(y)\frac{\partial y}{\partial x} - \cos(yz)(y\frac{\partial z}{\partial x} + z\frac{\partial y}{\partial x}) = 0$$

   $$3x^2 - \operatorname{sen}(yx)(\frac{\partial y}{\partial x}x + y) + \cos(z)\frac{\partial z}{\partial x} = 0$$

   evaluando en el punto $(0, 0, 0, 0)$ se tiene que $\partial y/\partial x(0, 0) = 0$.

2. En este caso se tiene que $F(0,\pi,0,0)=(0,0)$ y

$$D_{(x,y)}F(x,y,w,z)=\begin{pmatrix} 2x & \cos(y)-\operatorname{sen}(yz)z \\ 3x^2 & -\operatorname{sen}(yx)x \end{pmatrix}$$

por lo tanto

$$D_{(x,y)}F(0,\pi,0,0)=\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}$$

esta última no es una matriz invertible por lo que el teorema no es aplicable. Si se quisiera despejar (x,w) en función de (y,z) debemos observar la matriz

$$D_{(x,w)}F(x,y,z,w)=\begin{pmatrix} 2x & -3w^2 \\ 3x^2 & -2w \end{pmatrix}$$

que tampoco es invertible en los puntos $(0,0,0,0)$ y $(0,\pi,0,0)$, por lo que nuevamente el teorema no es aplicable.

□

Demostración. (Teorema de la función implícita)
Definamos la función

$$\begin{aligned} f:\Omega\subseteq\mathbb{R}^{n+m} &\to \mathbb{R}^{n+m} \\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto f(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{x},F(\boldsymbol{x},\boldsymbol{y})) \end{aligned}$$

que es evidentemente una función de clase $\mathcal{C}^1$ gracias a que F lo es. Su Jacobiano en el punto $(\boldsymbol{a},\boldsymbol{b})$ es

$$Df(\boldsymbol{a},\boldsymbol{b})=\begin{pmatrix} I_{m\times m} & 0 \\ D_xF(\boldsymbol{a},\boldsymbol{b}) & D_yF(\boldsymbol{a},\boldsymbol{b}) \end{pmatrix}$$

que es invertible ya que por hipótesis $D_yF(\boldsymbol{a},\boldsymbol{b})$ lo es. Además $f(\boldsymbol{a},\boldsymbol{b})=(\boldsymbol{a},0)$. Podemos entonces aplicar el teorema de la función inversa (teorema 5.2) a la función f. Existen abiertos $U,V\subseteq\mathbb{R}^{n+m}$ tales que $(\boldsymbol{a},\boldsymbol{b})\in U$, $(\boldsymbol{a},\boldsymbol{0})\in V$ y $f:U\to V$ es biyectiva con inversa de clase $\mathcal{C}^1$. Definamos ahora el conjunto $W=\{\boldsymbol{z}\in\mathbb{R}^m/(\boldsymbol{z},\boldsymbol{0})\in V\}$, entonces $\boldsymbol{a}\in W$ y W es un conjunto abierto de $\mathbb{R}^m$ pues V es abierto (demostrar que W es abierto queda de *tarea*). Luego, para cada $\boldsymbol{z}\in W$ la ecuación $f(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{z},\boldsymbol{0})$ tiene un único par $(\boldsymbol{x}_z,\boldsymbol{y}_z)$ como solución, es decir,

$$(\boldsymbol{x}_z,F(\boldsymbol{x}_z,\boldsymbol{x}_z))=(\boldsymbol{z},\boldsymbol{0})$$

esto último implica que $\boldsymbol{x}_z=\boldsymbol{z}$ y por lo tanto $F(\boldsymbol{z},\boldsymbol{y})=\boldsymbol{0}$. Como $\boldsymbol{y}_z$ es único dado $\boldsymbol{z}$, esto define una función $G(\boldsymbol{z})=\boldsymbol{y}_z$. Evidentemente $G(\boldsymbol{a})=\boldsymbol{b}$, y

$$(\boldsymbol{z},\boldsymbol{y}_z)=f^{-1}(\boldsymbol{z},\boldsymbol{0})=(\boldsymbol{z},G(\boldsymbol{z}))$$

De esta forma, G es la función que estamos buscando y como f^{-1} es de clase $\mathcal{C}^1$, entonces la ecuación anterior dice que G es de clase $\mathcal{C}^1$ también. Se tiene que G satisface la ecuación $F(\boldsymbol{z},G(\boldsymbol{z}))=0$, y como F y G son de clase $\mathcal{C}^1$ podemos calular al Jacobiano de G usando esta ecuación y la regla de la cadena

$$D_xF(\boldsymbol{z},G(\boldsymbol{z}))I_{m\times m}+D_yF(\boldsymbol{z},G(\boldsymbol{z}))DG(\boldsymbol{z})=0$$

de donde se obtiene el resultado

$$DG(\boldsymbol{z})=-[D_yF(\boldsymbol{z},G(\boldsymbol{z}))]^{-1}D_xF(\boldsymbol{z},G(\boldsymbol{z}))$$

para todo $\boldsymbol{z}\in W$. ■

5.4. Ejercicios

Contracciones y teorema del punto fijo

Ejercicio 1. Sea T una función tal que T^k es contractante. Demuestre que T tiene un único punto fijo.

Ejercicio 2. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ una función contractante en $B(\mathbf{0}, \delta)$, con constante L conocida. Demuestre que si $\|f(\mathbf{0})\| < \delta(1-L)$ entonces existe un único punto fijo de f en dicha vecindad.

Ejercicio 3. Muestre que la ecuación integral

$$u(x) = \frac{1}{2}\int_0^1 e^{-x-y}\cos(u(y))dy$$

tiene una y sólo una solución en el espacio $C([0,1],\mathbb{R})$. Para ello defina el operador

$$F(u)(x) = \frac{1}{2}\int_0^1 e^{-x-y}\cos(u(y))dy$$

y demuestre que es contractante. En $C([0,1],\mathbb{R})$ use la norma del supremo.

Ejercicio 4. Considere la ecuación integral

$$u(x) = 5 + \int_0^x \operatorname{sen}(u(s)+x)ds$$

Muestre que la ecuación posse una y solo una solución en $C([0,1/2],\mathbb{R})$

Ejercicio 5. Determinar para qué valores de a y b la función

$$T(x,y) = (a\cos(x+y), b\ln(1+x^2+y^2))$$

es una contración.

Ejercicio 6. Programe en su calculadora o en Matlab un algoritmo que encuentre la solución del sistema del ejemplo anterior. Para ello defina $x_0 = 0, y_0 = 0$, y genere una sucesión dada por la siguiente recurrencia:

$$\begin{aligned} x_{n+1} &= (1/2)\cos(x_n + y_n) \\ y_{n+1} &= (1/3)\ln(1+x^2+y^2) + 5 \end{aligned}$$

Ejercicio 7. Sean $g_i : \mathbb{R}^n \to \mathbb{R}$ funciones de clase $\mathcal{C}^1(\mathbb{R})$, $\forall i \in \{1,2,\dots n\}$. Suponga además que

$$\text{máx}\{\|\nabla g_i(\boldsymbol{x})\|_\infty\} < \frac{1}{2n} \qquad \forall i \in \{1,2,\dots n\}$$

Pruebe que el sistema $x_1 = g_1(\boldsymbol{x}), x_2 = g_2(\boldsymbol{x}), \dots x_n = g_n(\boldsymbol{x})$ tiene una única solución en $\mathbb{R}^n$, donde $\boldsymbol{x} = \{x_1, x_2, \dots x_n\}$.

Indicación. Recuerde que el espacio $(\mathbb{R}^n, \|\cdot\|_1)$ es de Banach.

Teorema de la función inversa

Ejercicio 8. Sea $f : U \to \mathbb{R}^n$ una función de clase $\mathcal{C}^1$ en el abierto $U \subset \mathbb{R}^n$. Sea $\boldsymbol{a} \in U$ tal que $f'(\boldsymbol{a})$ es invertible. Muestre que

$$\lim_{r \to 0} \frac{\operatorname{vol}(f(B(\boldsymbol{a}, r))}{\operatorname{vol}(B(\boldsymbol{a}, r)} = |\det f'(\boldsymbol{a})|$$

Ejercicio 9. Como en el ejercicio anterior muestre que si $f'(\boldsymbol{a})$ no es invertible entonces

$$\lim_{r \to 0} \frac{\operatorname{vol}(f(B(\boldsymbol{a}, r))}{\operatorname{vol}(B(\boldsymbol{a}, r)} = 0$$

Ejercicio 10. Para

$$f(u, v) = (u + [\log(v)]^2, uw, w^2)$$

1. Muestre que f no es inyectiva
2. Encuentre un dominio Ω de manera que la función sea inyectiva en Ω
3. Calcule $D(f^{-1})(1, 0, 1)$ cuando f se considera en Ω.

Ejercicio 11. Sea $f(x, y) = (x^2 - y^2, 2xy)$

1. Demuestre que para todo $(a, b) \neq (0, 0)$, f es invertible localmente en (a, b).
2. Demostrar que f no es inyectiva.
3. Calcular aproximadamente $f^{-1}(-3{,}01; 3{,}98)$. Use la fórmula de Taylor. Note que $f(1, 2) = (-3, 4)$

Ejercicio 12. Sean $f(u, v) = (u^2 + u^2 v + 10v, u + v^3)$

1. Encuentre el conjunto de puntos en los cuales f es invertible localmente.
2. Compruebe que $(1, 1)$ pertenece al conjunto obtenido en la parte 1), encontrar (aproximadamente) el valor de $f^{-1}(11{,}8, 2{,}2)$.

Teorema de la función implícita

Ejercicio 13. Si $f(\frac{x}{z}, \frac{y}{z}) = 0$, Calcular

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y}$$

Ejercicio 14. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ tal que

$$f(xy, z - 2x) = 0 \tag{*}$$

Suponga que existe $z(x_0, y_0) = z_0$ tal que $f(x_0 y_0, z_0 - 2x_0) = 0$ y que para todo (x, y) en una vecindad de (x_0, y_0), los puntos $(x, y, z(x, y))$ cumplen la ecuación (*). Encuentre que condiciones debe satisfacer f para que se cumpla lo anterior y demuestre que dad la condición anterior la función $z(x, y)$ cumple la igualdad

$$x_0 \frac{\partial z}{\partial x}(x_0, y_0) - y_0 \frac{\partial z}{\partial y}(x_0, y_0) = 2x_0$$

Ejercicio 15. La función z está definida por la siguiente ecuación

$$f(x - az, y - bz) = 0$$

donde F es una función diferenciable. Encuentre a que es igual la expresión

$$a\frac{\partial z}{\partial x} + b\frac{\partial z}{\partial y}$$

Ejercicio 16. Sea $z = f(x, y)$ definida implícitamente por la ecuación

$$x^2 z e^y + y^2 e^x + y = 1$$

Si $g(u, v) = (u^2 + v + 1, v^2)$, calcule $\frac{\partial^2 f \circ g}{\partial u \partial v}(0, 0)$. Justifique.

Ejercicio 17. Considere el siguiente sistema no-lineal:

$$\begin{aligned} x_1^2 x_2 + x_2^2 x_1 + y_2^2 y_1 + y_2 &= 0 \\ x_1 x_2 y_1 + x_2 y_1 y_2 + y_2 y_1 x_1 + y_2 x_1 x_2 &= 0 \end{aligned}$$

1. Demuestre que es posible despejar (y_1, y_2) en función de (x_1, x_2) en torno a algún punto.
2. Encuentre los valores de :
$$\frac{\partial y_1}{\partial x_2}(1, -1) \; ; \; \frac{\partial y_2}{\partial x_2}(1, -1)$$

Ejercicio 18. Sea $z(x, y)$ una funcion diferenciable definida implicitamente por la ecuacion:

$$z = x \cdot f\left(\frac{y}{z}\right)$$

donde $f : \mathbb{R} \to \mathbb{R}$ es ua funcion derivable. Demuestre que

$$2(x, y) \cdot \nabla z(x, y) = z(x, y)$$

Indicación. Considere $F : \mathbb{R}^3 \to \mathbb{R}$ definida como $F(x, y, z) = z - x \cdot f(\frac{y}{z})$

Ejercicio 19. Considere el siguiente sistema no lineal:

$$\begin{aligned} \operatorname{sen}(x_1 \cdot y_1) + x_3 \cos(y_2) &= 0 \\ y_2 + x_1^2 + (\operatorname{senh}(x_2))^2 &= 0 \end{aligned}$$

Encuentre los valores de $\frac{\partial y_1}{\partial x_2}(1, 0, 0)$ y $\frac{\partial y_2}{\partial x_3}(1, 0, 0)$.

Indicación. Considere la función

$$\begin{array}{rcl} g: \mathbb{R}^3 \times \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x,y) & \mapsto & (u(x,y), v(x,y)) \end{array}$$

con $x = (x_1, x_2, x_3), y = (y_1, y_2)$, $u, v : \mathbb{R}^3 \times \mathbb{R}^2 \to \mathbb{R}$ son tales que

$$u(x,y) = \operatorname{sen}(x_1 \cdot y_1) + x_3 \cos(y_2)$$

$$v(x,y) = y_2 + x_1^2 + (\operatorname{senh}(x_2))^2$$

Ejercicio 20. Mostrar que cerca del punto $(x, y, u, v) = (1, 1, 1, 1)$ podemos resolver el sistema

$$\begin{array}{rcl} xu + yvu^2 & = & 2 \\ xu^3 + y^2v^4 & = & 2 \end{array}$$

de manera única para u y v como funciones de x e y. Calcular $\frac{\partial u}{\partial x}$

Ejercicio 21. Considere el sistema

$$\begin{array}{rcl} \dfrac{x^4 + y^4}{x} & = & u \\ \operatorname{sen}(x) + \cos(y) & = & v \end{array}$$

Determine cerca de cuales puntos (x, y) podemos resolver x e y en términos de u y v

Ejercicio 22. Analizar la solubilidad del sistema

$$\begin{array}{rcl} 3x + 2y + z^2 + u + v^2 & = & 0 \\ 4x + 3y + zu^2 + v + w + 2 & = & 0 \\ x + z + w + u^2 + 2 & = & 0 \end{array}$$

para u, v, w en términos de x, y, z cerca de $x = y = z = 0$, $u = v = 0$ y $w = -2$.

Ejercicio 23.

1. Hallar $\frac{dy}{dx}|_{x=1}$ y $\frac{d^2y}{dx^2}|_{x=1}$ si
$$x^2 - 2xy + y^2 + x + y - 2 = 0$$

2. Hallar $\frac{dy}{dx}$ y $\frac{d^2y}{dx^2}$ si
$$\ln(\sqrt{x^2 + y^2}) = a \cdot \arctan\left(\frac{y}{x}\right) \quad (a \neq 0)$$

3. Hallar $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$ si
$$x\cos(y) + y\cos(z) + z\cos(x) = 1$$

4. Las funciones y y z de la variable independiente x se dan por el sistema de ecuaciones $xyz = a$ y $x + y + z = b$; Hallar $\frac{dy}{dx}, \frac{dz}{dx}, \frac{d^2y}{dx^2}, \frac{d^2z}{dx^2}$.

Ejercicio 24. Sea la funcion z dada por la ecuación

$$x^2 + y^2 + z^2 = \phi(ax + by + cz)$$

donde ϕ es una función cualquiera diferenciable y a, b, c son constantes; demostrar que:

$$(cy - bz)\frac{\partial z}{\partial x} + (az - cx)\frac{\partial z}{\partial y} = bx - ay$$

Ejercicio 25. Demostrar que la función z, determinada por la ecuación

$$y = x\phi(z) + \chi(z)$$

Satisface la ecuación

$$\frac{\partial^2 z}{\partial x^2}\left(\frac{\partial z}{\partial y}\right)^2 - 2\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y^2}\left(\frac{\partial z}{\partial x}\right)^2 = 0$$

Ejercicio 26. Sea $f(x, y, z) = 0$ tal que sus derivadas parciales son distintas de cero. Por lo tanto es posible escribir: $x = x(y, z)$, $y = y(x, z)$, $z = z(x, y)$, donde estas funciones son diferenciables. Muestre que

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1$$

Ejercicio 27. Muestre que las ecuaciones

$$\begin{aligned} x^2 - y^2 - u^3 + v^2 + 4 &= 0 \\ 2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0 \end{aligned}$$

determinan funciones $u(x, y)$ y $v(x, y)$ en torno al punto $x = 2, y = -1$ tales que $u(2, -1) = 2$ y $v(2, -1) = 1$. Calcular $\frac{\partial u}{\partial x}$ y $\frac{\partial v}{\partial y}$

CAPÍTULO 6

Complementos de Cálculo Diferencial

Nos interesa extender algunas reglas de derivación a casos más generales que frecuentemente se utilizan en la ingeniería y daremos los detalles del teorema del cambio de variables en integración.

6.1. Reglas de derivación adicionales

Teorema 6.1. (Regla de Leibniz de derivación)
Sea $f : \mathbb{R}^2 \to \mathbb{R}$ de clase $\mathcal{C}^1$. $\alpha, \beta : \mathbb{R} \to \mathbb{R}$ funciones diferenciables. Entonces la función $F : \mathbb{R} \to \mathbb{R}$ definida por

$$F(x) = \int_{\alpha(x)}^{\beta(x)} f(x,t)dt$$

es diferenciable y su derivada es

$$\frac{d}{dx}\int_{\alpha(x)}^{\beta(x)} f(x,t)dt = f(x,\beta(x))\beta'(x) - f(x,\alpha(x))\alpha'(x) + \int_{\alpha(x)}^{\beta(x)} \frac{\partial}{\partial x} f(x,t)dt$$

Demostración. Definamos la función

$$G(z,x) = \int_0^z f(x,t)dt$$

por el 1^{er} teorema fundamental del cálculo (teorema 1.8) sabemos que

$$\frac{\partial}{\partial z} G(z,x) = f(x,z)$$

demostremos que

$$\frac{\partial}{\partial x} G(z,x) = \int_0^z \frac{\partial}{\partial x} f(x,t)dt$$

para ello notemos que

$$\begin{aligned} & G(z,x+h)-G(z,x)-h\int_0^z \frac{\partial}{\partial x}f(x,t)dt \\ = & \int_0^z f(x+h,t)+f(x,t)-h\frac{\partial}{\partial x}f(x,t)dt \\ = & \int_0^z\int_0^1 (\frac{\partial}{\partial x}f(x+yh,t)-\frac{\partial}{\partial x}f(x,t))hdydt \end{aligned}$$

La función $\partial f/\partial x(\cdot,\cdot)$ es continua, y por lo tanto, uniformemente continua sobre $[x-|h|,x+|h|]\times[0,z]$, asi, dado $\varepsilon>0$, existe $\delta>0$ tal que si $\|(x,y)-(\bar{x},\bar{y})\|<\delta$ entonces $|\partial f/\partial x(x,y)-\partial f/\partial x(\bar{x},\bar{y})|<\varepsilon/z$. Entonces si $|h|<\delta$ se tiene que

$$|\frac{\partial}{\partial x}f(x+yh,t)-\frac{\partial}{\partial x}f(x,t)|<\varepsilon\ \ \forall t\in[0,z]\ \forall y\in[0,1]$$

lo que implica que

$$|G(z,x+h)-G(z,x)-h\int_0^z\frac{\partial}{\partial x}f(x,t)dt|<\varepsilon|h|$$

si $|h|<\delta$ de donde se sigue el resultado. Luego

$$\int_{\alpha(x)}^{\beta(x)} f(x,t)dt = G(\beta(x),x)-G(\alpha(x),x)$$

y aplicando el resultado anterior se tiene que

$$\begin{aligned} & \frac{d}{dx}\int_{\alpha(x)}^{\beta(x)} f(x,t)dt \\ = & \frac{\partial}{\partial z}G(\beta(x),x)\beta'(x)+\frac{\partial}{\partial x}G(\beta(x),x)\frac{dx}{dx}-\frac{\partial}{\partial z}G(\alpha(x),x)\alpha'(x)-... \\ & ...-\frac{\partial}{\partial x}G(\beta(x),x)\frac{dx}{dx} \\ = & f(x,\beta(x))\beta'(x)-f(\alpha(x),x)\alpha'(x)+\int_0^{\beta(x)}\frac{\partial f}{\partial x}(x,t)dt-\int_0^{\alpha(x)}\frac{\partial f}{\partial x}(x,t)dt \\ = & f(x,\beta(x))\beta'(x)-f(x,\alpha(x))\alpha'(x)+\int_{\alpha(x)}^{\beta(x)}\frac{\partial}{\partial x}f(x,t)dt \end{aligned}$$

■

Teorema 6.2. *Sean $A\subseteq\mathbb{R}^n$ y $B\subseteq\mathbb{R}^m$ abiertos y $f:A\times B\to\mathbb{R}$ de clase $\mathcal{C}^1$. Sea $Q\subset B$ conjunto elemental cerrado. Entonces la función*

$$F(\boldsymbol{x})=\int_Q f(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{y}$$

es de clase $\mathcal{C}^1$ y

$$\frac{\partial}{\partial x_i}F(\boldsymbol{x})=\int_Q\frac{\partial}{\partial x_i}f(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{y}$$

para todo $i\in\{1,...,n\}$

Demostración. Sea $\boldsymbol{x}_0 \in A$

$$
\begin{aligned}
& F(\boldsymbol{x}_0+\boldsymbol{h})-F(\boldsymbol{x}_0)-\sum_{i=1}^{n} h_i \int_Q \frac{\partial}{\partial x_i} f(\boldsymbol{x}_0,\boldsymbol{y}) d\boldsymbol{y} \\
= & \int_Q f(\boldsymbol{x}_0+\boldsymbol{h},\boldsymbol{y})-f(\boldsymbol{x}_0,\boldsymbol{y})-\sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0,\boldsymbol{y}) d\boldsymbol{y} \\
= & \int_Q (f(\boldsymbol{x}_0+\boldsymbol{h},\boldsymbol{y})-f(\boldsymbol{x}_0,\boldsymbol{y})-\nabla_x f(\boldsymbol{x}_0,\boldsymbol{y})\cdot\boldsymbol{h}) d\boldsymbol{y} \\
= & \int_Q \int_0^1 \left[(\nabla_x f(\boldsymbol{x}_0+t\boldsymbol{h},\boldsymbol{y})-\nabla_x f(\boldsymbol{x}_0,\boldsymbol{y}))\cdot\boldsymbol{h}\right] dt d\boldsymbol{y}
\end{aligned}
$$

La función $\nabla_x f(\cdot,\cdot)$ es continua en $\bar{B}(\boldsymbol{x}_0,1)\times Q$ que es compacto, por cual que es uniformemente continua. Luego dado $\varepsilon>0$ existe $\delta>0$ tal que si $\|(\boldsymbol{x},\boldsymbol{y})-(\boldsymbol{x}^*,\boldsymbol{y}^*)\|<\delta \Rightarrow \|\nabla_x f(\boldsymbol{x},\boldsymbol{y})-\nabla_x f(\boldsymbol{x}^*,\boldsymbol{y})\|<\varepsilon/\mathrm{vol}(Q)$. entonces si $\|\boldsymbol{h}\|<\delta$ se tiene que

$$
\|\nabla_x f(\boldsymbol{x}_0+t\boldsymbol{h},\boldsymbol{y})-\nabla_x f(\boldsymbol{x}_0,\boldsymbol{y})\| < \frac{\varepsilon}{\mathrm{vol}(Q)} \ \forall\, t\in[0,1]\ \forall\, \boldsymbol{y}\in Q
$$

lo que implica que

$$
\begin{aligned}
& |F(\boldsymbol{x}_0+\boldsymbol{h})-F(\boldsymbol{x}_0)-\sum_{i=1}^{n} h_i \int_Q \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0,\boldsymbol{y}) d\boldsymbol{y}| \\
\leq & \int_Q \int_0^1 \|\nabla f(\boldsymbol{x}_0+t\boldsymbol{h},\boldsymbol{y})-\nabla f(\boldsymbol{x}_0,\boldsymbol{y})\|\|\boldsymbol{h}\| dt d\boldsymbol{y} \\
\leq & \int_Q \int_0^1 \frac{\varepsilon}{\mathrm{vol}(Q)}\|\boldsymbol{h}\| = \varepsilon\|\boldsymbol{h}\|
\end{aligned}
$$

de este modo

$$
\left|F(\boldsymbol{x}_0+\boldsymbol{h})-F(\boldsymbol{x}_0)-\sum_{i=1}^{n} h_i \int_Q \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0,\boldsymbol{y}) d\boldsymbol{y}\right| \leq \varepsilon\|\boldsymbol{h}\|
$$

siempre que $\|\boldsymbol{h}\|<\delta$, es decir, F es diferenciable es $\boldsymbol{x}_0$ y

$$
\frac{\partial}{\partial x_i} F(\boldsymbol{x}_0) = \int_Q \frac{\partial}{\partial x_i} f(\boldsymbol{x}_0,\boldsymbol{y}) d\boldsymbol{y}
$$

para todo $i\in\{1,...,n\}$. La continuidad de las derivadas parciales queda de ejercicio. ■

6.2. La fórmula de cambio de variables

Recordemos la fórmula de cambio de variables

$$
\int_{f(\Omega)} g(\boldsymbol{x}) d\boldsymbol{x} = \int_\Omega g(f(\boldsymbol{y}))|\det Df(\boldsymbol{y})| d\boldsymbol{y}
$$

Para dar sentido a la fórmula se requiere

- $f: U \to V$ biyectiva de clase $\mathcal{C}^1$ y con inversa de clase $\mathcal{C}^1$. (Esta clase de funciones recibe el nombre de difeomorfismo)

- Ω es un compacto, cuya frontera es la unión finita de grafos de funciones continuas
- $g : f(\Omega) \to \mathbb{R}$ es continua.

En estas circunstancias las funciones

$$g^*(\boldsymbol{x}) = \begin{cases} 0 & \boldsymbol{x} \notin f(\Omega) \\ g(\boldsymbol{x}) & \boldsymbol{x} \in f(\Omega) \end{cases}$$

y

$$(g \circ f)^*(\boldsymbol{y})|\det Df(\boldsymbol{y})| = \begin{cases} 0 & \boldsymbol{y} \notin \Omega \\ g(f(\boldsymbol{y}))|\det Df(\boldsymbol{y})| & \boldsymbol{y} \in \Omega \end{cases}$$

son integrables.
Para demostrar la fórmula de cambio de variables, tenemos que dar una definición de integrabilidad un poco más fuerte.

Definición 6.1. $A \subseteq \mathbb{R}^n$ se dice v-medible si A es compacto y su frontera es la unión finita de grafos de funciones continuas.

Ejemplo 6.1. $A = [a, b]^n$ con a y b finitos es v-medible

Ejemplo 6.2. Si A es v-medible, y f es un difeomorfismo, entonces $f(A)$ en v-medible

Observamos que si A es v-medible, entonces podemos definir el volumen de A como

$$\text{vol}(A) = \int 1_A(\boldsymbol{x})d\boldsymbol{x}$$

donde

$$1_A(\boldsymbol{x}) = \begin{cases} 0 & \boldsymbol{x} \notin A \\ 1 & \boldsymbol{x} \in A \end{cases}$$

Definición 6.2. Sea A un conjunto v-medible. Una descomposición $\mathcal{D}$ de A es una colección de conjuntos $C_1, ..., C_k$ tales que

- C_i es v-medible, $i = 1, ..., k$
- $A = C_1 \cup ... \cup C_k$
- $int(C_i \cap C_j) = \phi$, si $i \neq j$

Se define también el diametro de la descomposición $\mathcal{D}$ como

$$\text{diam}(\mathcal{D}) = \max_{1 \leq i \leq k} \text{diam}(C_i)$$

donde el diametro de un conjunto A v-medible es igual a $\text{diam}(A) = \sup\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{x}, \boldsymbol{y} \in A\}$

Dada la descomposición $\mathcal{D}$ de A, consideremos el vector $\boldsymbol{\xi} = (\xi_1, ..., \xi_k)$ tal que $\xi_i \in C_i$ $\forall i \in \{1, \ldots, k\}$

Definición 6.3. Sea $f : A \to \mathbb{R}$ una función acotada. Definimos la suma de Riemann asociada a $(\mathcal{D}, \boldsymbol{\xi})$ como

$$S(f, \mathcal{D}, \boldsymbol{\xi}) = \sum_{i=1}^{k} f(\xi_i)\text{vol}(C_i)$$

Definición 6.4. Decimos que f es v-integrable sobre A, conjunto v-medible, si existe $I \in \mathbb{R}$ tal que

$$\forall \varepsilon > 0\, \exists \delta > 0 \text{ tal que } \forall \mathcal{D} \text{ desc. de } A,$$

$$(\operatorname{diam}(\mathcal{D}) < \delta\, \wedge \xi \text{ asociado a } \mathcal{D}) \Rightarrow |S(f, \mathcal{D}, \boldsymbol{\xi}) - I| < \varepsilon$$

Si tal I existe, es único y denotamos $I = \int_A f(\boldsymbol{x}) d\boldsymbol{x}$. La demostración de esto último queda de *tarea.*

Con esta definición podemos seguir paso a paso la demostración ya hecha para probar la proposicion siguiente

Proposición 6.1. $f : A \to \mathbb{R}$ continua $\Rightarrow$ f es v-integrable

También se puede demostrar que si $f : A \to \mathbb{R}$ es acotada y continua salvo sobre el grafo de un número finito de funciones continuas, entonces f es v-integrable.

Teorema 6.3. (Teorema del cambio de variables)
Sea Ω v-medible y $f : U \to V$ un difeomorfismo, $\Omega \subseteq U$. Sea $g : f(\Omega) \to \mathbb{R}$ v-integrable. Entonces $g \circ f : \Omega \to \mathbb{R}$ es v-integrable y

$$\int_{f(\Omega)} g(\boldsymbol{x}) d\boldsymbol{x} = \int_{\Omega} g(f(\boldsymbol{y})) |\det Df(\boldsymbol{y})| d\boldsymbol{y}$$

La demostración del teorema tiene dos partes: Primero el caso en que f es una función lineal y luego el caso general.

Demostración. (Caso f lineal)
La demostración ya fue hecha para $\mathbb{R}^3$, y se extiende de manera análoga a $\mathbb{R}^n$, por lo que supondremos cierto el resultado en el caso lineal. Asi tenemos, si T es lineal que

$$\operatorname{vol}(T(A)) = \int_{T(A)} 1 = \int_A 1 |\det T| = \operatorname{vol}(A) |\det T| \ \forall A \ \ v\text{-medible}$$

■

Para el caso general necesitamos dos lemas previos que se describen a continuación

Lema 6.1. *Sea U abierto, $X \subset U$ compacto y $\varphi : U \times U \to \mathbb{R}$ continua tal que $\varphi(\boldsymbol{x}, \boldsymbol{x}) = 1 \forall \boldsymbol{x} \in U$. Entonces $\forall \varepsilon > 0\, \exists \delta > 0$ tal que*

$$\boldsymbol{x}, \boldsymbol{y} \in X \ \|\boldsymbol{x} - \boldsymbol{y}\| < \delta \Rightarrow |\varphi(\boldsymbol{x}, \boldsymbol{y}) - 1| < \varepsilon$$

Demostración. Puesto que X es compacto, entonces $X \times X \subset U \times U$ también es compacto, por lo que φ será uniformemente continua sobre $X \times X$. Luego, dado $\varepsilon > 0\, \exists \delta > 0$ tal que $\|\boldsymbol{x} - \boldsymbol{z}\| + \|\boldsymbol{y} - \boldsymbol{w}\| = \|(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{z}, \boldsymbol{w})\| < \delta \Rightarrow |\varphi(\boldsymbol{x}, \boldsymbol{y}) - \varphi(\boldsymbol{z}, \boldsymbol{w})| < \varepsilon$, en particular, si $\|\boldsymbol{x} - \boldsymbol{y}\| < \delta$ entonces $\|(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{x}, \boldsymbol{x})\| < \delta$ lo que implica que $|\varphi(\boldsymbol{x}, \boldsymbol{y}) - \varphi(\boldsymbol{x}, \boldsymbol{x})| = |\varphi(\boldsymbol{x}, \boldsymbol{y}) - 1| < \varepsilon$ ■

Lema 6.2. *Sean $U, V \subseteq \mathbb{R}^n$ abiertos y $f : U \to V$ un difeomorfismo de clase $\mathcal{C}^1$. Sea $X \subset U$ un compacto v-medible y $M = \sup\{\|Df(\boldsymbol{x})\| : \boldsymbol{x} \in X\}$ ($\|\cdot\|$ es cualquier norma matricial, por ejemplo $\|A\| = \max_{1 \le i \le n}\{\sum_{j=1}^{n} |a_{ij}|\}$). Entonces,*

$$\operatorname{vol}(f(X)) \le M^n \operatorname{vol}(X)$$

Demostración. Demostremos primero el caso en que X es un cubo, con centro p y lado $2a$, es decir,

$$X = [p_1 - a, p_1 + a] \times ... \times [p_n - a, p_n + a] = \prod_{i=1}^{n} [p_i - a, p_i + a]$$

Por la desigualdad del valor medio tenemos que $|f_i(x) - f_i(p)| \leq Ma$ y por lo tanto

$$\|f(\boldsymbol{x}) - f(p)\|_\infty \leq Ma \ \forall \boldsymbol{x} \in X$$

es decir, $f(\boldsymbol{x})$ está a distancia a lo más Ma de $f(p)$ para todo $\boldsymbol{x} \in X$, entonces

$$\begin{aligned} f(X) &\subseteq [f_1(p) - Ma, f_1(p) + Ma] \times \ldots \times [f_n(p) - Ma, f_n(p) + Ma] \\ &= \prod_{i=1}^{n} [f_i(p) - Ma, f_i(p) + Ma] \end{aligned}$$

o sea que $\text{vol}(f(X)) \leq \text{vol}(\prod_{i=1}^{n} [f_i(p) - Ma, f_i(p) + Ma]) = M^n (2a)^n = M^n \text{vol}(X)$

El caso general lo hacemos por aproximación. Sea $\varepsilon > 0$, entonces existe un abierto θ tal que $X \subset \theta \subset U$ y $\|Df(\boldsymbol{x})\| \leq M + \varepsilon$ para todo $\boldsymbol{x} \in \theta$ (Por la continuidad de Df).
El conjunto X se puede cubrir por un número finito de cubos con interiores disjuntos contenidos en θ

$$X \subset \bigcup_{i=1}^{k} C_i$$

con $int(C_i) \cap int(C_j) = \phi$ si $i \neq j$. Además los cubos se pueden suponer tan pequeños que

$$\sum_{i=1}^{k} \text{vol}(C_i) \leq \text{vol}(X) + \varepsilon$$

Luego

$$f(X) \subseteq \bigcup_{i=1}^{k} f(C_i) \Rightarrow \text{vol}(f(X)) \leq \sum_{i=1}^{k} \text{vol}(f(C_i)) \leq \sum_{i=1}^{k} M_i^n \text{vol}(C_i)$$

donde $M_i = \sup\{\|Df(\boldsymbol{x})\| : \boldsymbol{x} \in C_i\} \leq M + \varepsilon$ por lo tanto $\text{vol}(f(X)) \leq (M + \varepsilon)^n (\text{vol}(X) + \varepsilon)$.
Como esta desigualdad es válida para todo $\varepsilon > 0$ concluimos que

$$\text{vol}(f(X)) \leq M^n \text{vol}(X)$$

■

Demostración. (Caso general)
Para finalizar con la demostración en el caso general, consideremos una descomposición $\mathcal{D} = \{C_1, ..., C_k\}$ de Ω y puntos $\xi_i \in C_i$ para $i : \{i, \ldots, k\}$. Entonces los conjuntos $f(C_i)$ definen una descomposición de $f(\Omega)$ y los puntos $f(\xi_i) \in f(C_i)$. A esta descomposición la denotamos por $\mathcal{D}f$.
Usando la desigualdad del valor medio se puede demostrar que existen constantes a_1, a_2 tales que

$$a_1 \cdot \text{diam}(\mathcal{D}f) \leq \text{diam}(\mathcal{D}) \leq a_2 \cdot \text{diam}(\mathcal{D}f)$$

gracias a que Ω es compacto. Definamos $T_i = f'(\xi_i) \in M_{n \times n}$ para $i \in \{1, \ldots, k\}$, y

$$N_i = \sup\{\|T_i^{-1} f'(\boldsymbol{x})\| : \boldsymbol{x} \in C_i\} \ M_i = \sup\{\|T_i (f^{-1})'(\boldsymbol{y})\| : \boldsymbol{y} \in f(C_i)\}$$

con $\|\cdot\|$ la misma del lema anterior. Entonces se tiene que

$$\begin{aligned} \mathrm{vol}(f(C_i)) &= \mathrm{vol}(T_iT_i^{-1}f(C_i)) \\ &= |\det(T_i)|\mathrm{vol}(T_i^{-1}f(C_i)) \\ &\leq |\det(T_i)|N_i^n\mathrm{vol}(C_i) \end{aligned}$$

de igual manera

$$\mathrm{vol}(C_i) \leq |\det(T_i^{-1})|\mathrm{vol}(f(C_i))M_i^n$$

De lo anterior se obtiene que

$$\begin{aligned} \mathrm{vol}(C_i)|\det(T_i)| - \mathrm{vol}(f(C_i)) &\leq \mathrm{vol}(f(C_i))(M_i^n - 1) \\ \mathrm{vol}(C_i)|\det(T_i)| - \mathrm{vol}(f(C_i)) &\geq \mathrm{vol}(C_i)|\det(T_i)|(1 - N_i^n) \end{aligned}$$

Definamos también $\varphi(\boldsymbol{x}, \boldsymbol{y}) = \|(f'(\boldsymbol{y}))^{-1}f'(\boldsymbol{x})\|$, de este modo $\varphi(\boldsymbol{x}, \boldsymbol{x}) = 1$ para todo $\boldsymbol{x} \in \Omega$. Luego, gracias a los lemas anteriores, dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $\mathrm{diam}(\mathcal{D}) < \delta$ entonces

$$|N_i^n - 1| < \varepsilon \ \wedge \ |M_i^n - 1| < \varepsilon$$

De todo lo anterior, se concluye que existe una constante B tal que

$$|\mathrm{vol}(C_i)|\det(T_i)| - \mathrm{vol}(f(C_i))| < B\varepsilon\mathrm{vol}(f(C_i))$$

Según la definición de v-integrable debemos comparar la suma de Riemann de $h = g \circ f|\det Df|$ asociada la partición $\mathcal{D}$ y a los puntos $\boldsymbol{\xi}$ con el supuesto valor de la integral: $I = \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}$

$$\begin{aligned} & \left|S(h, \mathcal{D}, \boldsymbol{\xi}) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| \\ = & \left|\sum_{i=1}^k g(f(\xi_i))|\det(T_i)|\mathrm{vol}(C_i) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| \\ \leq & \left|\sum_{i=1}^k g(f(\xi_i))|\det(T_i)|\mathrm{vol}(C_i) - \sum_{i=1}^k g(f(\xi_i))\mathrm{vol}(f(C_i))\right| + \dots \\ & + \left|\sum_{i=1}^k g(f(\xi_i))\mathrm{vol}(f(C_i)) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| \\ \leq & B\varepsilon\sum_{i=1}^k |g(f(\xi_i))|\mathrm{vol}(f(C_i)) + \dots \\ & + \left|\sum_{i=1}^k g(f(\xi_i))\mathrm{vol}(f(C_i)) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| \\ \leq & B\varepsilon\sum_{i=1}^k |g(f(\xi_i))|\mathrm{vol}(f(C_i)) + \left|S(g, \mathcal{D}f, f(\boldsymbol{\xi})) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| \end{aligned}$$

En la última desigualdad, gracias a que g es v-integrable sobre $f(\Omega)$ tenemos que la sumatoria de la izquierda está acotada y el término de la derecha se puede hacer tan pequeño como se desee tomando un δ suficientemente pequeño. Por lo tanto dado $\bar{\varepsilon} > 0$ existe $\bar{\delta} > 0$ tal que si $\mathrm{diam}\mathcal{D} < \bar{\delta}$, entonces

$$\left|S(h, \mathcal{D}, \boldsymbol{\xi}) - \int_{f(\Omega)} g(\boldsymbol{x})d\boldsymbol{x}\right| < \bar{\varepsilon}$$

es decir, $g \circ f |\det Df|$ es v-integrable en Ω y

$$\int_{\Omega} g \circ f(\boldsymbol{y}) |\det Df(\boldsymbol{y})| d\boldsymbol{y} = \int_{f(\Omega)} g(\boldsymbol{x}) d\boldsymbol{x}$$

pues la integral es única. ■

6.3. Ejercicios

Reglas de derivación adicionales

Ejercicio 1. Sea $f(x,y,z,t) = \int\limits_{e^{x+y+z}}^{\ell(1+|x+y|)} (xt + log_z(t) + z)dt$. Calcular $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial t}$ en aquellos puntos donde existan.

Ejercicio 2. Considere la siguiente ecuación integral

$$u(x) = 5 + \int_0^x \operatorname{sen}(u(s) + x)ds$$

En el capítulo 5 se demostró que tiene una única solución en $C([0,1/2],\mathbb{R})$.
Derivando dos veces la ecuación integral (justifique por que se puede derivar), encuentre la ecuación diferencial que satisface su solución

Ejercicio 3. Considere las funciones

$$F(x) = \int_0^{g(x)} f(x+t, xt)dt, \quad G(x,y) = \int_0^{y^2} g(x,y,z)dz$$

Calcule $F'(x)$ y $\frac{\partial G}{\partial y}(x,y)$

Ejercicio 4. Considere la función

$$\begin{aligned} u(x,t) &= \frac{1}{2}[f(x+ct) + f(x-ct)] + \frac{1}{2c}\int\limits_{x-ct}^{x+ct} g(s)ds \\ &+ \frac{1}{2c}\int\limits_0^t \int\limits_{x-c(t-s)}^{x+c(t-s)} F(\sigma,s)d\sigma ds \end{aligned}$$

Muestre que

$$\begin{aligned} \frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} &= F(x,t) \\ u(x,0) &= f(x) \\ \frac{\partial u}{\partial t}(x,0) &= g(x) \end{aligned}$$

Ejercicio 5. Se desea resolver la ecuación de ondas

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c}\frac{\partial^2 \phi}{\partial t^2} = 0$$

donde c es una constante distinta de 0.
Haga el cambio de variables $\xi = x + ct$ y $\eta = x - ct$ de modo que se obtenga

$$\phi(x,t) = \psi(\xi(x,t), \eta(x,t))$$

donde $\psi(\xi,\eta)$ es la incógnita. En base a este cambio de variables demuestre que:

$$\frac{\partial^2\phi}{\partial x^2}-\frac{1}{c}\frac{\partial^2\phi}{\partial t^2}=4\frac{\partial^2\psi}{\partial\eta\partial\xi}(\xi(x,y),\eta(x,y))$$

y en base a este resultado demuestre que toda solución ϕ de clase $\mathcal{C}^2$ de la ecuación de ondas se escribe

$$\phi(x,t)=f(\xi)+g(\eta)$$

con f y g funciones de variable real.

Cambio de variables

Ejercicio 6. Sea B_n la bola unitaria en $\mathbb{R}^n$. Muestre que

$$Vol(B_4)=2\int_{B_3}\sqrt{1-(x^2+y^2+z^2)}dxdydz.$$

Tomando coordenadas esféricas muestre que el volumen de B_4 es igual a

$$2\int_{-\pi/2}^{\pi/2}\int_0^{2\pi}\int_0^1 r^2\operatorname{sen}(\phi)\sqrt{1-r^2}drd\theta d\phi$$

y concluya que $Vol(B_4)=\pi^2/2$. En general muestre que el volumen de la bola de radio r es $r^4\pi^2/2$.

Ejercicio 7. Sea $f:\mathbb{R}^n\to\mathbb{R}^n$ un difeomorfismo de clase $\mathcal{C}^1$ (es decir, un homeomorfismo con f y f^{-1} de clase $\mathcal{C}^1$), que satisface que $f(B)\subset B$, donde B es la bola unitaria cerrada y $|\det f'(x)|<1$ para todo $x\in B$. Demuestre que para toda función continua $g:B\to\mathbb{R}^n$ se tiene que

$$\lim_{k\to\infty}\int_{f^k(B)}g(x)dx=0$$

donde f^k es f compuesta consigo misma k veces.

Ejercicio 8. Calcular $\int\limits_S\int dxdy$ donde S es el dominio limitado por

$$\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)^2=\frac{x^2}{h^2}-\frac{y^2}{k^2}$$

Indicación. Use $x=\operatorname{arc}\cos(\varphi)$ e $y=\operatorname{sen}(\varphi)$.

CAPÍTULO 7

Optimización no Lineal

Veremos en forma general los teoremas de los multiplicadores de Lagrange y el teorema de Karush-Kuhn-Tucker los cuales, bajo determinadas condiciones, nos dan condiciones necesarias para la existencia de máximos o mínimos de una función sujeta a varias restricciones.

7.1. Problema general de optimización

Recordemos algunos conceptos ya vistos. Si tenemos un problema de la forma

$$\begin{array}{lll} P) & \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ & & \boldsymbol{x} \in S \end{array} \tag{7.1}$$

Donde $f : \mathbb{R}^n \to \mathbb{R}$ y $S \subset \mathbb{R}^n$. Si $S = \mathbb{R}^n$ tenemos el caso de optimización sin restricciones y en tal caso si $\boldsymbol{x}_0$ es un mínimo local

$$f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}) \ \forall \boldsymbol{x} \in B(\boldsymbol{x}_0, \delta)$$

por lo tanto para cada $\boldsymbol{d} \in \mathbb{R}^n$ y $t \approx 0$ se tiene que

$$f(\boldsymbol{x}_0 + t\boldsymbol{d}) - f(\boldsymbol{x}_0) \geq 0$$

Si f es diferenciable, tras dividir por t y aplicar límite cuando $t \to 0^+$ se obtiene

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0$$

como $\boldsymbol{d}$ es arbitrario la condición de primer orden para el caso irrestricto es

$$\nabla f(\boldsymbol{x}_0) = 0$$

En el caso en que S sea una parte de $\mathbb{R}^n$ no necesariamente $\boldsymbol{x}_0 + t\boldsymbol{d} \in S$, por ejemplo si $S = \mathbb{R}^n_+$ y $\boldsymbol{x}_0$ se encuentra en la frontera de S se tiene un caso en que no necesariamente $\boldsymbol{x}_0 + t\boldsymbol{d} \in S$ y cualquier $\boldsymbol{d}$ arbitrario no nos sirve para concluir la condición de primer orden.

Definición 7.1. (Espacio tangente)
El espacio tangente a S en $\boldsymbol{x}_0$ corresponde al conjunto

$$T_S(\boldsymbol{x}_0) = \{\boldsymbol{d} \in \mathbb{R}^n : \exists t_n \to 0^+,\ \boldsymbol{d}_n \to \boldsymbol{d} \text{ tal que } \boldsymbol{x}_0 + t_n\boldsymbol{d}_n \in S\}$$

y es un conjunto tal que $\boldsymbol{d} = 0 \ \in T_S(\boldsymbol{x}_0)$ y en caso de que $\boldsymbol{x}_0 \in \text{int}(S)$ entonces $T_S(\boldsymbol{x}_0) = \mathbb{R}^n$.

Teorema 7.1. *Si $\boldsymbol{x}_0$ es un mínimo local de f en S, entonces se cumple la condición necesaria*

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0 \ \forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0)$$

Demostración. Escogamos $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ en S y $\{t_n\}_{n\in\mathbb{N}}$ en $\mathbb{R}$ tales que $\boldsymbol{x}_n \to \boldsymbol{x}_0$ y $t_n \to 0^+$ en la medida que $n \to \infty$, por lo que $\boldsymbol{x}_0 + t_n\boldsymbol{d}_n \in S$. Si $\boldsymbol{x}_0$ es un mínimo local de f en S entonces

$$f(\boldsymbol{x}_0 + t_n\boldsymbol{d}_n) - f(\boldsymbol{x}_0) \geq 0$$

dividiendo por t_n y aplicando límite cuando $n \to \infty$ se tiene

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0$$

■

Teorema 7.2. *Si $\boldsymbol{x}_0 \in S$ y se tiene la condición suficiente*

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} > 0 \ \forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$$

entonces $\boldsymbol{x}_0$ es un mínimo local estricto de f en S.

Demostración. Si $\boldsymbol{x}_0$ no es mínimo local de f en S entonces existe $\boldsymbol{x}_n \in S$ tal que $\boldsymbol{x}_n \to \boldsymbol{x}_0$ y $\boldsymbol{x}_n \neq \boldsymbol{x}_0$ de forma tal que

$$f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0)$$

Podemos definir

$$\boldsymbol{d}_n = \frac{\boldsymbol{x}_n - \boldsymbol{x}_0}{\|\boldsymbol{x}_n - \boldsymbol{x}_0\|}$$

tal que $\|\boldsymbol{d}_n\| = 1$ y entonces $\{\boldsymbol{d}_n\}$ es acotada por lo que tiene una subsucesión convergente. En base a esto es posible suponer que converge a $\boldsymbol{d}_0 \in T_S(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$. La expansión de Taylor de $f(\boldsymbol{x}_n)$ está dada por

$$f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\| R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|)$$

esto más la condición $f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0)$ conducen a

$$\nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\| R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \leq 0$$

diviendo por $\|\boldsymbol{x}_n - \boldsymbol{x}_0\|$ y tomando límite cuando $n \to \infty$

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d}_0 \leq 0$$

lo cual contradice la hipótesis. ■

Nota 7.1. La condición necesaria no es suficiente y la condición suficiente no es necesaria.

Ejemplo 7.1. Consideremos los casos:

1. $f(x) = x^3$ y $S = [-1, 1]$. La condición necesaria nos dice que el mínimo local es $x_0 = 0$, entonces $f'(x_0) = 0$ y $T_S(x_0) = \mathbb{R}$. Se tiene que x_0 no es mínimo local y la condición necesaria se cumple pero no asegura la minimalidad local.

2. $f(x) = x^2$ y $S = [-1, 1]$. La condición necesaria nos dice que el mínimo local es $x_0 = 0$, entonces $f'(x_0) = 0$ y $T_S(x_0) = \mathbb{R}$. Se tiene que x_0 es mínimo local y la condición suficiente no se cumple pese a que $x_0 = 0$ cumple con ser un mínimo local del problema.

Definición 7.2. (Espacio normal)
El espacio normal a S en $\boldsymbol{x}_0$ corresponde al conjunto

$$N_S(\boldsymbol{x}_0) = \{\boldsymbol{u} \in \mathbb{R}^n : \boldsymbol{u} \cdot \boldsymbol{d} \leq 0 \ \forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0)\}$$

y a partir de esta definición la condición necesaria corresponde a

$$-\nabla f(\boldsymbol{x}_0) \in N_S(\boldsymbol{x}_0)$$

Teorema 7.3. *Sea $f : D \subset \mathbb{R}^n \to \mathbb{R}$ una función convexa y diferenciable en un convexo D tal que $S \subset D$. Entonces f tiene un mínimo global en $\boldsymbol{x}_0$ en S si y sólo si*

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0 \ \forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0)$$

Demostración. Una forma de demostrar esta propiedad es mediante la condición necesaria la cual implica que f tiene un mínimo global en $\boldsymbol{x}_0$. Sea $\boldsymbol{x} \in S$, entonces si definimos $\boldsymbol{x}_\lambda = (1-t_k)\boldsymbol{x}_0 + t_k\boldsymbol{x}$ con $t_k \to 0$ se tiene que $\boldsymbol{x}_\lambda \in S$ que implica $\boldsymbol{x}_\lambda \to \boldsymbol{x}_0$ y obtenemos que $\boldsymbol{x} - \boldsymbol{x}_0 \ \in T_S(\boldsymbol{x}_0)$. A partir de la convexidad de f tenemos que

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)$$

esto más la condición necesaria implican

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$$

y se concluye que $\boldsymbol{x}_0$ es mínimo global de f en S. ∎

Los teoremas de esta sección corresponden a un tratamiento muy abstracto que provee condiciones generales, las cuales no son facilmente aplicables. Dicho esto, conviene dar más estructura al conjunto S y para aquello se puede plantear el problema (7.1) de la siguiente forma:

$$\begin{array}{lll} P) & \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ & \text{s.a} & g_i(\boldsymbol{x}) = 0 \quad \forall i \in \{1, \ldots, k\} \\ & & h_j(\boldsymbol{x}) \leq 0 \quad \forall j \in \{1, \ldots, m\} \end{array} \tag{7.2}$$

Resulta conveniente suponer que el número de restricciones es finito y en este caso el conjunto de restricciones es

$$S = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) = 0 \ \forall i \in I, \ h_j(\boldsymbol{x}) \leq 0 \ \forall j \in J\}$$

donde $I = \{1, \ldots, k\}$, $J = \{1, \ldots, m\}$ y $g_i, \ h_j : \mathbb{R}^n \to \mathbb{R}$ son funciones diferenciables.

Definición 7.3. Un vector $\boldsymbol{x}_0 \in \mathbb{R}^n$ es factible si $\forall i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, m\}$ se cumple que $g_i(\boldsymbol{x}_0) = 0$ y $h_j(\boldsymbol{x}_0) \leq 0$.

Definición 7.4. En el contexto de un problema de minimización diremos que un vector $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^k \times \mathbb{R}^m_+$ corresponde a una familia de multiplicadores de Karush-Kuhn-Tucker asociados con $\boldsymbol{x}_0 \in S$ si

$$\begin{aligned} \mu_j &\geq 0 \; \forall j \in J \\ \mu_j h_j(\boldsymbol{x}_0) &= 0 \; \forall j \in J \\ L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) &\leq L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \; \forall \boldsymbol{x} \in S \end{aligned}$$

donde

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \sum_{i \in I} \lambda_i g_i(\boldsymbol{x}) + \sum_{j \in J} \mu_j h_j(\boldsymbol{x})$$

es la función lagrangeano y es casi idéntica a la que presentamos en el capítulo 4.

Cuando aparezcan solo restricciones de igualdad diremos que $\boldsymbol{\lambda} \in \mathbb{R}^k$ corresponde a una familia de multiplicadores de Lagrange asociados con $\boldsymbol{x}_0 \in S$ si

$$L(\boldsymbol{x}_0, \boldsymbol{\lambda}) \leq L(\boldsymbol{x}, \boldsymbol{\lambda}) \; \forall \boldsymbol{x} \in S$$

donde las componentes de $\boldsymbol{\lambda}$ no tienen restricción de signo.

Definición 7.5. Sea $J(\boldsymbol{x}_0) = \{j : 1 \leq j \leq m \, , h_j(\boldsymbol{x}_0) = 0\}$ diremos que $J(\boldsymbol{x}_0)$ contiene p elementos $\{1, \ldots, p\}$ con $p \leq m$.

A partir de $J(\boldsymbol{x}_0)$ diremos que $\boldsymbol{x}_0 \in \mathbb{R}^n$ es regular si el conjunto

$$L = \{\nabla g_i(\boldsymbol{x}_0), \nabla h_j(\boldsymbol{x}_0) : i \in I, j \in J(\boldsymbol{x}_0)\}$$

es linealmente independiente.

7.2. Teorema de los Multiplicadores de Lagrange

Motivación: Por ahora nos restringiremos al caso de un problema de maximización o minimización que consta únicamente de restricciones de igualdad para extendernos de manera gradual a casos más generales. De hecho, rara vez un sistema no tiene restricciones y por esto nos interesa ampliar lo que ya hemos visto sobre extremos restringidos.

La pregunta básica que queremos responder en esta sección, es ¿Cuándo una ecuación como $g(x, y, z) = 0$ define una superficie?

Los dos ejemplos siguientes son muy elocuentes:

1. $x^2 + y^2 + z^2 - 1 = 0$
2. $x(z - x^2 + y^2) = 0$

$x^2 + y^2 + z^2 - 1 = 0$ ciertamente es una superficie (es la superficie de la esfera de radio 1 y centro en el origen). Sin embargo $x(z - x^2 + y^2) = 0$ no es una superficie.
Este ejemplo nos muestra que la respuesta es local. El problema con la superficie 2. ocurre en el punto $(0, 0, 0)$ donde $\nabla g(0, 0, 0) = (0, 0, 0)$.

Definición 7.6. En general, si el punto (x_0, y_0, z_0) es tal que $g(x_0, y_0, z_0) = 0$ y $\nabla g(x_0, y_0, z_0) \neq 0$ entonces la ecuación $g(x, y, z) = 0$ define una superficie, al menos en una vecindad de (x_0, y_0, z_0); en efecto, si $\nabla g(x_0, y_0, z_0) \neq 0$, entonces alguna derivada parcial es no nula, supongamos sin perdida de generalidad que

$$\frac{\partial g}{\partial y}(x_0, y_0, z_0) \neq 0$$

entonces, por el teorema de la función implícita, es posible despejar y en función de (x, z), es decir, podemos escribir $y = y(x, z)$, y por lo tanto el conjunto $S = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$ es localmente el grafo de una función, es decir, es una superficie.

7.2.1. Condiciones de primer orden para extremos restringidos

Considerando la definición 2.10, en todo lo que sigue en esta sección los teoremas que presentamos para maximización (minimización) son válidos para la minimización (maximización), esto se debe a que un criterio de maximización sobre una función f cóncava (convexa) es análogo a la minimización de $-f$ que es una función convexa (cóncava). La única salvedad es que cuando aparezcan desigualdades sobre los resultados algunas de estas se invierten (esto se indicará en los casos que ocurre).

Definición 7.7. (Espacios normal y tangente, caso particular)
Sean $g : \mathbb{R}^3 \to \mathbb{R}$ una función diferenciable, $(x_0, y_0, z_0) \in \mathbb{R}^3$ tal que $g(x_0, y_0, z_0) = 0$, y supongamos que $\nabla g(x_0, y_0, z_0) \neq 0$. Llamemos S a la superficie que define la ecuación $g(x, y, z) = 0$ en torno a (x_0, y_0, z_0). Se define el espacio normal a S en el punto (x_0, y_0, z_0) como el espacio vectorial generado por el gradiente de g en el punto (x_0, y_0, z_0) y se denota $N_s(x_0, y_0, z_0)$.

$$N_s(x_0, y_0, z_0) = \langle \nabla g(x_0, y_0, z_0) \rangle$$

El espacio tangente a S en el punto (x_0, y_0, z_0) se denota $T_s(x_0, y_0, z_0)$ y se define como el ortogonal del espacio normal

$$T_s(x_0, y_0, z_0) = N_s(x_0, y_0, z_0)^{\perp} = \{(x, y, z) \in \mathbb{R}^3 : (x, y, z) \cdot \nabla g(x_0, y_0, z_0) = 0\}$$

Nota 7.2. Los espacios normal y tangente, son espacios vectoriales, y no pasan necesariamente por el punto (x_0, y_0, z_0).

Con la definición anterior, si $\sigma : A \subseteq \mathbb{R} \to \mathbb{R}^3$ es una función tal que $\sigma(t) \in S\ \forall t$, y $\sigma(0) = (x_0, y_0, z_0)$, entonces $\sigma'(0) \in T_s(x_0, y_0, z_0)$.
¿Es posible alcanzar cada $\boldsymbol{d} \in T_s(x_0, y_0, z_0)$ de esta forma, con una función σ adecuada?
Supongamos que $\nabla g(x_0, y_0, z_0) \neq 0$. Sea $\boldsymbol{d} \in T_s(x_0, y_0, z_0)$. Elijamos $\boldsymbol{b} \neq 0$ de manera tal que $\boldsymbol{b} \perp \langle \{v, \nabla g(x_0, y_0, z_0)\} \rangle$; siempre existe tal $\boldsymbol{b}$ pues el espacio vectorial $\langle \{\boldsymbol{d}, \nabla g(x_0, y_0, z_0)\} \rangle$ tiene dimensión 2. Consideremos el sistema de ecuaciones

$$\begin{aligned} g(x, y, z) &= 0 \\ (x - x_0, y - y_0, z - z_0) \cdot \boldsymbol{b} &= 0 \end{aligned}$$

el Jacobiano del sistema anterior en el punto (x_0, y_0, z_0) es

$$\begin{pmatrix} \partial g/\partial x(x_0, y_0, z_0) & \partial g/\partial y(x_0, y_0, z_0) & \partial g/\partial z(x_0, y_0, z_0) \\ b_x & b_y & b_z \end{pmatrix}$$

Como $\boldsymbol{b} \perp \nabla g(x_0, y_0, z_0)$, la matriz Jacobiana tiene rango 2, y por lo tanto siempre es posible escoger dos columnas tales que la matriz resultante sea invertible. El teorema de la función implícita nos asegura entonces que podremos siempre despejar dos variables en función de una. Supongamos sin perdida de generalidad que es posible despejar (y,z) en función de x, entonces

$$\begin{aligned} g(x,y(x),z(x)) &= 0 \\ (x-x_0, y(x)-y_0, z(x)-z_0)\cdot \boldsymbol{b} &= 0 \end{aligned}$$

Si se define $\sigma(t)=(t+x_0, y(t+x_0), z(t+z_0))$, entonces

$$g(\sigma(t))=0 \;\wedge\; (\sigma(t)-(x_0,y_0,z_0))\cdot \boldsymbol{b}=0$$

para todo t en una vecindad del cero. La función $\sigma(t)$ es de clase $\mathcal{C}^1$ pues las funciones $y(x), z(x)$ lo son, por lo tanto, derivando las ecuaciones anteriores con respecto a t y evaluando en $t=0$ se obtiene

$$\nabla g(x_0,y_0,z_0)\cdot \sigma'(0)=0 \;\wedge\; \sigma'(0)\cdot \boldsymbol{b}=0$$

entonces por la elección de $\boldsymbol{b}$ no queda otra posibilidad más que $\sigma'(0)\parallel \boldsymbol{d}$. Si se define $\bar{\sigma}(t)=\sigma(\|\boldsymbol{d}\|t/\|\sigma'(0)\|)$ entonces $\bar{\sigma}'(0)=\boldsymbol{d}$ y $g(\bar{\sigma}(t))=0$ para todo t en una vecindad del cero. De esta forma concluimos que

$$T_s(x_0,y_0,z_0)=\{\sigma'(0)/\sigma(t)\in S\ \forall t\ ,\ \sigma(0)=(x_0,y_0,z_0)\}$$

Teorema 7.4. (Teorema de los multiplicadores de Lagrange, caso particular)
Consideremos el siguiente problema de optimización

$$\begin{array}{lll} P) & \min\limits_{x,y,z} & f(x,y,z) \\ & s.a & g(x,y,z)=c \end{array}$$

con f,g funciones diferenciables. Si (x_0,y_0,z_0) es una solución del problema P), entonces existe $\lambda\in\mathbb{R}$ tal que

$$\nabla f(x_0,y_0,z_0)=\lambda\nabla g(x_0,y_0,z_0)$$

Demostración. Sea $\sigma: A\subseteq\mathbb{R}\to\mathbb{R}^3$ tal que $g(\sigma(t))=0$ y $\sigma(0)=(x_0,y_0,z_0)$, es decir, $\sigma(t)\in S$ para t en una vecindad del origen. Como (x_0,y_0,z_0) es un mínimo local de f restringido a S, entonces

$$\left.\frac{d}{dt}f(\sigma(t))\right|_{t=0}=\nabla f(x_0,y_0,z_0)\cdot\sigma'(0)=0$$

por lo hecho anteriormente, esto equivale a que $\nabla f(x_0,y_0,z_0)\cdot \boldsymbol{d}=0\ \forall v\in T_s(x_0,y_0,z_0)\Rightarrow \nabla f(x_0,y_0,z_0)\in T_s(x_0,y_0,z_0)^{\perp}=N_s(x_0,y_0,z_0)$, y entonces por definición de $N_s(x_0,y_0,z_0)$ existe $\lambda\in\mathbb{R}$ tal que

$$\nabla f(x_0,y_0,z_0)=\lambda\nabla g(x_0,y_0,z_0)$$

■

Definición 7.8. (Espacios normal y tangente, caso general)
Sean $g_i(\boldsymbol{x}):\mathbb{R}^n\to\mathbb{R}\ \forall i\in I$ una función diferenciable, $\boldsymbol{x}_0\in\mathbb{R}^n$ tal que $g_i(\boldsymbol{x}_0)=0\ \forall i\in I$, y supongamos que $\nabla g_i(\boldsymbol{x}_0)\neq 0$. Llamemos S a la superficie en $\mathbb{R}^n$ que define la ecuación $g_i(\boldsymbol{x})=0$ en torno a $\boldsymbol{x}_0$. Por analogía con el caso particular definimos el espacio normal a S en $\boldsymbol{x}_0$ como

$$N_s(\boldsymbol{x}_0)=\langle\{\nabla g_1(\boldsymbol{x}_0),...\nabla g_k(\boldsymbol{x}_0)\}\rangle$$

y el espacio tangente a S en $\boldsymbol{x}_0$ como

$$T_s(\boldsymbol{x}_0)=N_s(\boldsymbol{x}_0)^{\perp}=\{\boldsymbol{d}\in\mathbb{R}^n:\boldsymbol{d}\cdot\nabla g_i(\boldsymbol{x}_0)=0\ \forall i\in I\}$$

Lema 7.1. *$\forall \boldsymbol{d} \in T_s(\boldsymbol{x}_0)$ existe $\sigma : A \subseteq \mathbb{R} \to \mathbb{R}^n$, $0 \in A$ con $g_i(\sigma(t)) = 0 \, \forall t \in A$, $\sigma(t) \in S \, \forall t \in A$ y además $\sigma(0) = \boldsymbol{x}_0$ tal que $\sigma'(0) = \boldsymbol{d}$. Es decir*

$$T_s(\boldsymbol{x}_0) = \{\sigma'(0) : \sigma(t) \in S \, \forall t \ , \ \sigma(0) = \boldsymbol{x}_0\}$$

Demostración. Sean $\boldsymbol{d} \in T_s(\boldsymbol{x}_0)$ y $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-k-1}\}$ una base ortonormal del espacio

$$\langle \{\boldsymbol{d}, \nabla g_1(\boldsymbol{x}_0), ..., \nabla g_k(\boldsymbol{x}_0)\} \rangle^{\perp}$$

Consideremos el sistema de $n-1$ ecuaciones

$$\begin{array}{rcl} g_1(\boldsymbol{x}) & = & 0 \\ & \vdots & \\ (\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{b}_1 & = & 0 \\ & \vdots & \\ (\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{b}_{n-k-1} & = & 0 \end{array}$$

El vector $\boldsymbol{x}_0$ satisface las ecuaciones, y la matriz Jacobiana del sistema es

$$\begin{pmatrix} \nabla g_1(\boldsymbol{x}_0) \\ \vdots \\ \nabla g_k(\boldsymbol{x}_0) \\ \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_{n-k-1} \end{pmatrix}$$

que es una matriz de rango $n-1$, por lo que podemos seleccionar $n-1$ columnas tales que la matriz resultante sea invertible, y gracias al teorema de la función implícita (teorema 5.3) podemos despejar $n-1$ variables en función de la restante en una vecindad de $\boldsymbol{x}_0$. Entonces existe $\sigma : A \to \mathbb{R}^n$, con $A = (-\varepsilon, \varepsilon)$, definido de manera análoga al caso particular, tal que $g_i(\sigma(t)) = 0 \ \forall t \in A$ y $(\sigma(t) - \boldsymbol{x}_0) \cdot b_i = 0 \ \forall t, \ i \in \{1, ..., n-k-1\}$.

Derivando las ecuaciones anteriores con respecto a t y evaluando en $t = 0$ se obtiene que

$$\nabla g_i(\boldsymbol{x}_0) \cdot \sigma'(0) = 0 \ \forall i \in \{1, ..., k\} \ \wedge \ \sigma'(0) \cdot b_j = 0 \ \forall j \in \{1, ..., n-k-1\}$$

lo que implica que $\sigma'(0) \parallel \boldsymbol{d}$. Definiendo $\bar{\sigma}(t) = \sigma(\|\boldsymbol{d}\| t / \|\sigma'(0)\|)$ tenemos que $\bar{\sigma}'(0) = \boldsymbol{d}$, lo que nos da el lema, pues la otra inclusión $\{\sigma'(0) : \sigma(t) \in S \, \forall t, \, \sigma(0) = \boldsymbol{x}_0\} \subseteq T_s(\boldsymbol{x}_0)$ es directa (análoga a la demostración en el caso particular). ■

Teorema 7.5. (Teorema de los multiplicadores de Lagrange,caso general)
Consideremos el siguiente problema de optimización

$$\begin{array}{lll} P) & \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ & s.a & g_i(\boldsymbol{x}) = c \quad \forall i \in \{1, \ldots k\} \end{array} \tag{7.3}$$

donde $f, g_i : \mathbb{R}^n \to \mathbb{R}$ son funciones diferenciables. Supongamos que f alcanza un mínimo local en $\boldsymbol{x}_0 \in S$, donde $S = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) = 0\}$, y que $Dg(\boldsymbol{x}_0) \in M_{k \times n}$ tiene rango k. Entonces existe un vector $\boldsymbol{\lambda} \in \mathbb{R}^k$ tal que

$$\nabla f(\boldsymbol{x}_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0)$$

Puesto que $Dg(\boldsymbol{x}_0)$ es de rango k, el espacio $N_s(\boldsymbol{x}_0)$ es un espacio vectorial con $\dim N_s(\boldsymbol{x}_0) = k$, *y por lo tanto* $\dim T_s(\boldsymbol{x}_0) = n - k$.

Demostración. A partir del lema 7.1 tenemos que $\forall \sigma : A \subseteq \mathbb{R} \to \mathbb{R}^n$ tal que $g(\sigma(t)) = 0 \wedge \sigma(0) = \boldsymbol{x}_0$

$$\frac{d}{dt} f(\sigma(t))\bigg|_{t=0} = \nabla f(\boldsymbol{x}_0) \cdot \sigma'(0) = 0$$

puesto que $\boldsymbol{x}_0$ es un mínimo local de f restringido a S. Lo anterior es equivalente, gracias al lema, a que $\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall \boldsymbol{d} \in T_s(\boldsymbol{x}_0)$, o sea, $\nabla f(\boldsymbol{x}_0) \in N_s(\boldsymbol{x}_0)$ lo que implica que

$$\nabla f(\boldsymbol{x}_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0)$$

para ciertos $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, que es lo que se quería demostrar. ■

Los teoremas 7.4 y 7.5 dan una condición necesaria de optimalidad con restricciones de igualdad. También son válidos para un problema de maximización, esto porque de acuerdo a la definición de función convexa (definición 2.6) tenemos que minimizar una función f convexa equivale a maximizar la función $-f$ que resulta ser cóncava.

Sea $\boldsymbol{x}_0$ un mínimo de f restringida a S, se cumple que

$$\nabla f(\boldsymbol{x}_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0)$$

Para la maximización tendríamos lo siguiente

$$-\nabla f(\boldsymbol{x}_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0)$$

Esto sugiere que si escribimos el Lagrangeano para el caso de la minimización como

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{i=1}^{k} \lambda_i g_i(\boldsymbol{x})$$

Para un problema de maximización debe expresarse

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^{k} \lambda_i g_i(\boldsymbol{x})$$

En la práctica, la ventaja que genera el cambio de signo es que la interpretación de los multiplicadores de Lagrange es directa en problemas de maximización y minimización. Antes de dar la interpretación presentaremos algunos ejemplos que nos mostrarán esto en el camino para formalizar el resultado con el teorema de la envolvente. Al momento de resolver no hay diferencia para el caso en que trabajamos con restricciones de igualdad.

Nota 7.3. Ninguna de las condiciones del teorema puede relajarse y además el teorema no nos garantiza que los multiplicadores sean no nulos. Si sucede que, por ejemplo, la matriz que define $Dg(\boldsymbol{x}_0)$ no es de rango completo (no es de rango k) o los gradientes de la función objetivo y las restricciones son linealmente dependientes, el sistema que definen las condiciones de primer orden de la función Lagrangeano no tiene solución.

La condición de independencia lineal entre los gradientes de la función objetivo y las restricciones se tiene particularmente en el caso en que el número de restricciones es estrictamente menor que la dimensión del espacio sobre el cual trabajamos. De esto es que la hipótesis de independencia lineal resulta muy restrictiva.

Los dos ejemplos que siguen son ilustrativos de lo que pasa cuando no se cumplen las condiciones del teorema o al menos un multiplicador se anula.

Ejemplo 7.2. Consideremos el problema

$$\begin{array}{cc} \min\limits_{x,y} & x + y^2 \\ \text{s.a} & x - y^2 = 0 \end{array}$$

Observemos que $\nabla f(x,y)|_{(0,0)} = (1,0)$ y $\nabla g(x,y)|_{(0,0)} = (1,0)$. Entonces, los gradientes son linealmente dependientes y no se cumple que $\nabla f(x,y)|_{(0,0)} - \lambda \nabla g(x,y)|_{(0,0)} = (0,0)$ a menos que $\lambda = 1$ lo que hace que el sistema que definen las condiciones de primer orden del Lagrangeano no tenga solución dado que el par $(1,0)$ no cumple la restricción del problema.

Ejemplo 7.3. Consideremos el problema

$$\begin{array}{cc} \min\limits_{x,y} & x^2 + y^2 \\ \text{s.a} & x = 0 \end{array}$$

Observemos que $\nabla f(x,y)|_{(0,0)} = (0,0)$ y $\nabla g(x,y)|_{(0,0)} = (1,0)$. Entonces no se cumple que $\nabla f(x,y)|_{(0,0)} - \lambda \nabla g(x,y)|_{(0,0)} = (0,0)$ a menos que $\lambda = 0$ lo cual nos lleva a que el sistema que definen las condiciones de primer orden del Lagrangeano no tenga solución.

7.2.2. Condiciones de segundo orden para extremos restringidos

Definición 7.9. Para el caso de minimización definiremos el conjunto de direcciones críticas como

$$K(\boldsymbol{x}_0) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I,\ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0\}$$

Mientras que para el caso de maximización definiremos el conjunto de direcciones críticas como

$$K(\boldsymbol{x}_0) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I,\ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \leq 0\}$$

Ya vimos un teorema que nos da las condiciones de 2^{do} orden (teorema 2.17) el cual lo podemos enunciar, con una hipótesis menos estricta, de la siguiente forma

Teorema 7.6. *Sea $\boldsymbol{x}_0$ un mínimo local del problema (7.3) y supongamos que existe $\boldsymbol{\lambda} \in \mathbb{R}^k$ tal que*

$$\nabla f(\boldsymbol{x}_0) + \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0) = 0 \ \forall i \in I$$

considerando que $\{\nabla g_1(\boldsymbol{x}_0), \ldots, \nabla g_k(\boldsymbol{x}_0)\}$ es linealmente independiente. Entonces $\forall \boldsymbol{d} \in K(\boldsymbol{x}_0)$ se cumple que la forma cuadrática $\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}$ es semi definida positiva, es decir

$$\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d} \geq 0$$

Demostración. Sean $f, g_i \ \forall i \in I$ funciones de clase $\mathcal{C}^2$. Definamos $\sigma : A \subseteq \mathbb{R} \to \mathbb{R}^n$ tal que $\sigma(0) = \boldsymbol{x}_0$ y $g(\sigma(t)) = 0$, es decir $\sigma(t) \in S$. Entonces por la convexidad de f

$$\left.\frac{d^2}{dt^2} f(\sigma(t))\right|_{t=0} \geq 0$$

y por definición

$$\frac{d^2}{dt^2} f(\sigma(t))\Big|_{t=0} = \boldsymbol{d}_1^T(Hf(\boldsymbol{x}_0))\boldsymbol{d}_1 + \nabla f(\boldsymbol{x}_0)\cdot \boldsymbol{d}_2 \geq 0 \qquad (*)$$

Como $L(\boldsymbol{x}_0, \boldsymbol{\lambda}) = f(\boldsymbol{x}_0) + \sum_{i=1}^k \lambda_i g_i(\boldsymbol{x}_0)$ tenemos que

$$\nabla f(\boldsymbol{x}_0) + \boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0) = 0$$

donde $\boldsymbol{\lambda} \in \mathbb{R}^k$ y $g(\boldsymbol{x}_0) = (g_1(\boldsymbol{x}_0), \ldots, g_k(\boldsymbol{x}_0))$. Consideremos que

$$\boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0) = \boldsymbol{\lambda}^T \nabla g(\sigma(t)) = 0$$

y diferenciando con respecto a t se obtiene

$$\boldsymbol{d}_1^T(\boldsymbol{\lambda}^T Hg(\boldsymbol{x}_0))\boldsymbol{d}_1 + \boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0)\cdot \boldsymbol{d}_2 = 0 \qquad (**)$$

Sumando (**) a (*)

$$\boldsymbol{d}_1^T(Hf(\boldsymbol{x}_0) + \boldsymbol{\lambda}^T Hg(\boldsymbol{x}_0))\boldsymbol{d}_1 + \underbrace{(\nabla f(\boldsymbol{x}_0) + \boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0))}_{0}\cdot \boldsymbol{d}_2 \geq 0$$

$$\boldsymbol{d}_1^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}_1 \geq 0$$

Como $\boldsymbol{d}_1$ es arbitrario en $K(\boldsymbol{x}_0)$ se tiene que el resultado es válido para cualquier $\boldsymbol{d} \in K(\boldsymbol{x}_0)$ y llegamos a

$$\boldsymbol{d}(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d} \geq 0,\ \ \boldsymbol{d} \in K(\boldsymbol{x})$$

■

Teorema 7.7. *Sea $\boldsymbol{x}_0$ un máximo local de un problema de maximización con la estructura del problema (7.3) y supongamos que existe $\boldsymbol{\lambda} \in \mathbb{R}^k$ tal que*

$$\nabla f(\boldsymbol{x}_0) - \sum_{i=1}^k \lambda_i \nabla g_i(\boldsymbol{x}_0) = 0\ \forall i \in I$$

considerando que $\{\nabla g_1(\boldsymbol{x}_0), \ldots, \nabla g_k(\boldsymbol{x}_0)\}$ es linealmente independiente. Entonces $\forall \boldsymbol{d} \in K(\boldsymbol{x}_0)$ se cumple que la forma cuadrática $\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}$ es semi definida negativa, es decir

$$\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d} \leq 0$$

Demostración. Es análoga a la del teorema 7.6. Basta con tomar f cóncava y como $-f$ es convexa se concluye. ■

Ahora formularemos el teorema 2.17 tal como se vio en el capítulo 4.

Teorema 7.8. *Sea $\boldsymbol{x}_0$ un vector factible del problema (7.3) y supongamos que existe $\boldsymbol{\lambda} \in \mathbb{R}^k$ tal que*

$$\nabla f(\boldsymbol{x}_0) + \sum_{i=1}^k \lambda_i \nabla g_i(\boldsymbol{x}_0) = 0\ \forall i \in I$$

Para todo $\boldsymbol{d} \in K(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$ la forma cuadrática $\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}$ es definida positiva, es decir

$$\boldsymbol{d}^T(H_x L(\boldsymbol{x}, \boldsymbol{\lambda}))\boldsymbol{d} > 0$$

Entonces $\boldsymbol{x}_0$ es un mínimo local estricto.

Demostración. Consideremos que f es estrictamente convexa y supongamos que $\boldsymbol{x}_0$ no es un mínimo estricto de f restringida a S. Escogamos $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ en S tal que $\boldsymbol{x}_n \to \boldsymbol{x}_0$ y $f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0)$. Sea $\boldsymbol{x}_n = \boldsymbol{x}_0 + \gamma_n d_n$ con $\gamma_n > 0$ y

$$\boldsymbol{d}_n = \frac{\boldsymbol{x}_n - \boldsymbol{x}_0}{\|\boldsymbol{x}_n - \boldsymbol{x}_0\|}, \ \|\boldsymbol{d}_n\| = 1$$

Se tendrá que $\gamma_n = \|\boldsymbol{x}_n - \boldsymbol{x}_0\|$, $\gamma_n \to 0$ y $\{\boldsymbol{d}_n\}$ es acotada por lo que tiene una subsucesión convergente a $\boldsymbol{d}_0$.

Como $\{\boldsymbol{x}_n\} \in S$ se tiene que $g_i(\boldsymbol{x}_n) = 0 \ \forall i \in I$. Definamos $g(\boldsymbol{x}_0) = (g_1(\boldsymbol{x}_0), \ldots, g_k(\boldsymbol{x}_0))$ y de esta forma

$$g(\boldsymbol{x}_n) - g(\boldsymbol{x}_0) = 0 \ \Rightarrow \ g(\boldsymbol{x}_0 + \gamma_n \boldsymbol{d}_n) - g(\boldsymbol{x}_0) = 0$$

si dividimos esto por γ_n, tal que $\gamma_n \to 0$, en la medida que $n \to \infty$ se obtiene $\nabla g(\boldsymbol{x}_0) \cdot \boldsymbol{d}_0 = 0$.

Usando el teorema de Taylor (teorema 2.2), para todo $i \in I$ se cumple

$$g_i(\boldsymbol{x}_n) - g_i(\boldsymbol{x}_0) = \gamma_n \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d}_n + \frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (H g_i(\boldsymbol{x}_a)) \boldsymbol{d}_n = 0 \tag{*}$$

y también se tiene que

$$f(\boldsymbol{x}_n) - f(\boldsymbol{x}_0) = \gamma_n \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d}_n + \frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (H f(\boldsymbol{x}_b)) \boldsymbol{d}_n \leq 0 \tag{**}$$

con $\boldsymbol{x}_a = \alpha \boldsymbol{x}_n + (1-\alpha)\boldsymbol{x}_0$, $\alpha \in]0,1[$ y $\boldsymbol{x}_b = \beta \boldsymbol{x}_n + (1-\beta)\boldsymbol{x}_0$, $\beta \in]0,1[$, es decir $\boldsymbol{x}_a$ y $\boldsymbol{x}_b$ son combinaciones convexas de $\boldsymbol{x}_0$ y $\boldsymbol{x}_n$.

Consideremos $\frac{1}{n}\|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2 \geq f(\boldsymbol{x}_n) - f(\boldsymbol{x}_0)$, entonces $\frac{1}{n}\|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2 \geq 0$. A partir de (**)

$$\gamma_n \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d}_n + \frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (H f(\boldsymbol{x}_b)) \boldsymbol{d}_n \leq \frac{1}{n}\|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2$$

Multiplicando (*) por λ_i y aplicando la sumatoria $\sum_{i=1}^{k}(\cdot)$ para agregar las k restricciones formamos lo siguiente

$$\boldsymbol{\lambda}^T(g(\boldsymbol{x}_n) - g(\boldsymbol{x}_0)) = \gamma_n \boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0) \cdot \boldsymbol{d}_n + \frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (\boldsymbol{\lambda}^T H g(\boldsymbol{x}_a)) \boldsymbol{d}_n = 0 \tag{***}$$

Sumando (***) a (**) se obtiene

$$\gamma_n \underbrace{(\nabla f(\boldsymbol{x}_0) + \boldsymbol{\lambda}^T \nabla g(\boldsymbol{x}_0))}_{0} \cdot \boldsymbol{d}_n + \frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (H f(\boldsymbol{x}_b) + \boldsymbol{\lambda}^T H g(\boldsymbol{x}_a)) \boldsymbol{d}_n \leq \frac{1}{n}\|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2$$

$$\frac{\gamma_n^2}{2} \boldsymbol{d}_n^T (H f(\boldsymbol{x}_b) + \boldsymbol{\lambda}^T H g(\boldsymbol{x}_a)) \boldsymbol{d}_n \leq \frac{1}{n}\|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2$$

como $\gamma_n = \|\boldsymbol{x}_n - \boldsymbol{x}_0\|$

$$\boldsymbol{d}_n^T (H f(\boldsymbol{x}_b) + \boldsymbol{\lambda}^T H g(\boldsymbol{x}_a)) \boldsymbol{d}_n \leq \frac{2}{n}$$

y tomando límite cuando $n \to \infty$

$$\boldsymbol{d}_0^T (H f(\boldsymbol{x}_0) + \boldsymbol{\lambda}^T H g(\boldsymbol{x}_0)) \boldsymbol{d}_0 \leq 0$$

Observemos que en esto último aparecen $\boldsymbol{x}_a$ y $\boldsymbol{x}_b$. Como tomamos límite cuando $n \to \infty$ y se tiene que $\boldsymbol{x}_n \to \boldsymbol{x}_0$, entonces $\boldsymbol{x}_a \to \boldsymbol{x}_0$ y $\boldsymbol{x}_b \to \boldsymbol{x}_0$.

De esta forma hemos llegado a que cuando $\boldsymbol{x}_0$ es un mínimo local estricto de f restringida a S no es posible que la forma cuadrática $\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}$ sea semi definida negativa y como $\boldsymbol{d}$ es arbitrario se concluye la demostración. ■

Teorema 7.9. *Sea $\boldsymbol{x}_0$ un vector factible de un problema de maximización con la estructura del problema (7.3) y supongamos que existe $\boldsymbol{\lambda} \in \mathbb{R}^k$ tal que*

$$\nabla f(\boldsymbol{x}_0) - \sum_{i=1}^{k} \lambda_i \nabla g_i(\boldsymbol{x}_0) = 0 \ \forall i \in I$$

Para todo $\boldsymbol{d} \in K(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$ la forma cuadrática $\boldsymbol{d}^T(H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d}$ es definida negativa, es decir

$$\boldsymbol{d}^T(H_x L(\boldsymbol{x}, \boldsymbol{\lambda}))\boldsymbol{d} < 0$$

Entonces $\boldsymbol{x}_0$ es un máximo local estricto.

Demostración. Es análoga a la del teorema 7.8. Basta con tomar f cóncava y algún $\boldsymbol{x}_0$ factible suponiendo que no es máximo, como $-f$ es convexa se concluye. ■

Nota 7.4. Los teoremas 7.6 y 7.7 nos dan una condición necesaria de segundo orden mientras que los teoremas 7.8 y 7.9 nos dan una condición suficiente.

Los teoremas 7.7 y 7.9 nos dan una condición suficiente de segundo orden. Podría creerse que la condición clave para ambos teoremas es que el conjunto $\{\nabla g_i(\boldsymbol{x}_0)\}$ sea linealmente independiente $\forall i \in \{1, \ldots, k\}$, sin embargo la condición es que se cumpla que $g_i(\boldsymbol{x}_0) = 0 \ \forall i \in I$. Basta con que el Hessiano del Lagrangreano solo con respecto a $\boldsymbol{x}$ sea definido positivo en el conjunto de direcciones críticas y no en cualquier dirección arbitraria.

Ejemplo 7.4. En el ejemplo (2.10), diga si el punto $(x_0, y_0) = (-\frac{1}{2}, \frac{1}{2})$ es un máximo o un mínimo.

Solución. Para esto, construyamos primeramente el Lagrangeano del problema:

$$L(x, y, \lambda) = x^2 + y^2 - \lambda(y - x - 1)$$

Recordemos que el multiplicador de lagrange era $\lambda = 1$, por tanto, el Hessiano del Lagrangeano queda:

$$H_x L(x_0, y_0, \lambda) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

el cual es definido positivo para todo d, en particular para las direcciones del conjunto de direcciones críticas y por tanto el punto $(x_0, y_0) = (-\frac{1}{2}, \frac{1}{2})$ es un mínimo. □

Ejemplo 7.5. Optimice el valor de la función $f(x, y) = x$ sobre el conjunto de puntos (x, y) tales que $x^2 + 2y^2 = 3$.

Solución. El Lagrangeano para este problema queda de la siguiente forma:

$$L(x, y, \lambda) = x - \lambda(x^2 + 2y^2 - 3)$$

a partir del cual, aplicando el sistema lagrangeano (2.4) se obtiene

$$\begin{cases} 1 = 2\lambda x_0 \\ 4 = \lambda y_0 \\ x_0^2 + 2y_0^2 = 3 \end{cases}$$

$$\begin{cases} 1 = 2\lambda x_0 \\ 0 = 4\lambda y_0 \\ x_0^2 + 2y_0^2 = 3 \end{cases} \quad \Rightarrow \quad \begin{matrix} 1 = 2\lambda x_0 \\ x_0^2 + 2y_0^2 = 3 \end{matrix}$$

y por tanto, los posibles candidatos a extremos (λ no puede ser cero) son:

$$(x_1, y_1, \lambda_1) = \left(\sqrt{3}, 0, \frac{1}{2\sqrt{3}}\right)$$
$$(x_2, y_2, \lambda_2) = \left(-\sqrt{3}, 0, -\frac{1}{2\sqrt{3}}\right)$$

por otro lado se tiene que:

$$H_x L(x, y, \lambda) = \begin{pmatrix} -2\lambda & 0 \\ 0 & -4\lambda \end{pmatrix}$$

es decir, para

$$(x_1, y_1, \lambda_1) = \left(\sqrt{3}, 0, \frac{1}{2\sqrt{3}}\right)$$

$H_x L(x_1, y_1, \lambda_1)$ es definida negativa y por tanto $(x_1, y_1) = (\sqrt{3}, 0)$ es un máximo local mientras que para

$$(x_2, y_2, \lambda_2) = \left(-\sqrt{3}, 0, -\frac{1}{2\sqrt{3}}\right)$$

$H_x L(x_2, y_2, \lambda_2)$ es definida positiva y por lo tanto $(x_2, y_2) = (-\sqrt{3}, 0)$ es un mínimo local. $\square$

Ejemplo 7.6. Consideremos el siguiente problema

$$\begin{array}{cl} \max\limits_{x,y,z} & xy + yz + xz \\ \text{s.a} & x + y + z = 1 \end{array}$$

Nos interesa saber si la solución (en caso de que exista) efectivamente corresponde a un máximo local estricto.

Solución. El Lagrangeano corresponde a

$$L(x, \lambda) = xy + yz + xz - \lambda(x + y + z - 1)$$

Aplicando el sistema Lagrangeano llegamos a lo siguiente

$$\left\{ \begin{array}{l} y_0 + z_0 - \lambda = 0 \\ x_0 + z_0 - \lambda = 0 \\ x_0 + y_0 - \lambda = 0 \\ x_0 + y_0 + z_0 = 1 \end{array} \right. \quad \Rightarrow \quad \begin{array}{l} x_0 = y_0 \\ y_0 = z_0 \\ x_0 + y_0 + z_0 = 1 \end{array} \quad \Rightarrow \quad \begin{array}{l} x_0 = 1/3 \\ y_0 = 1/3 \\ z_0 = 1/3 \end{array}$$

Entonces $f(x_0, y_0, z_0) = 1/3$ y el valor del multiplicador de Lagrange es $\lambda = 2/3$.

Para las condiciones de segundo orden tenemos

$$Hf(x_0) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad Hg(x_0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Entonces,

$$H_x L(x_0, \lambda) = H_x f(x_0) + \lambda^T Hg(x_0) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Podríamos usar el criterio de los menores principales para determinar si la matriz resultante es semi definida positiva o negativa. Usando esto se obtiene $|H_1| = 0$, $|H_2| = -1$ y $|H_3| = 2$ lo cual bajo tal criterio nos dice que la matriz no es semi definida negativa ni tampoco semi definida positiva.

Los teoremas 7.6, 7.7, 7.8 y 7.9 establecen una condición de segundo orden que pide que $H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda})$ sea semi definida positiva (caso minimización) o semi definida negativa (caso maximización) a lo menos en $K(\boldsymbol{x})$.

Definamos $\boldsymbol{d} = (x, y, z)$ y entonces

$$\boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}))\boldsymbol{d} = x(y+z) + y(x+z) + z(x+y) \qquad (*)$$

Tengamos presente que en $K(x)$ se cumple que $\nabla g(x) \cdot d = 0$, entonces $x + y + z = 0$ por lo que podemos reescribir (*) como

$$d^T (H_x L(x_0, \lambda))d = -(x^2 + y^2 + z^2) < 0$$

observemos que $-(x^2 + y^2 + z^2) \leq 0$ pero restringido a $K(\boldsymbol{x})$ se tiene la desigualdad estricta y entonces la solución encontrada es un máximo local estricto. $\square$

7.2.3. Ejemplos de Microeconomía en varias dimensiones

Ahora daremos algunos ejemplos en los que se debe resolver una función Lagrangeano que consta de $n + 1$ variables. Es importante revisar estos problemas con detalle puesto que entregan una formulación para casos generalizados.

Ejemplo 7.7. En la ciudad de Titirilquén todas las empresas minimizan sus costos de operación. Supondremos que en esta economía existen n factores productivos que tienen un costo unitario de w_i por unidad.

En lo que sigue, supondremos que existen m empresas y que la empresa representativa de la ciudad elabora un único producto y su estructura de producción se puede representar mediante la siguiente función:

$$f(\boldsymbol{x}) = A \prod_{i=1}^{n} x_i^{\alpha_i}$$

Donde los parámetros A, α_i son positivos. Las curvas de nivel de esta función, en un caso particular, corresponden al siguiente gráfico

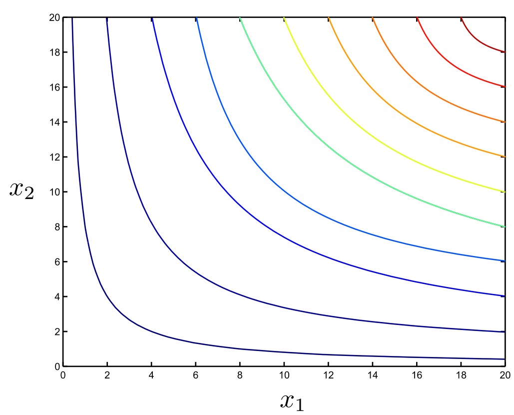


Figura 7.1: Función Cobb-Douglas $f(x_1, x_2) = x_1^{0,5} x_2^{0,5}$

Para fijar ideas, intuitivamente a partir del gráfico del caso particular, es posible deducir que el problema se puede resolver mediante Lagrangeano.

Lo que nos interesa obtener es la demanda por factores $x_i(\boldsymbol{w}, Q)$, donde x_i es la demanda por el insumo i-ésimo en función de $\boldsymbol{w}$ (vector de $\mathbb{R}^n_+$ que representa los costos de los n insumos) y de Q que representa un nivel de producción fijo. Esta demanda proviene del siguiente problema:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \sum_{i=1}^n w_i x_i \\ \text{s.a} & f(\boldsymbol{x}) \geq Q,\ x_i \geq 0 \end{array}$$

el cual se puede reescribir como

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \sum_{i=1}^n w_i x_i \\ \text{s.a} & f(\boldsymbol{x}) = Q \end{array}$$

De lo anterior obtendremos:

1. El planteamiento del problema de mínimo costo y el Lagrangeano del problema.
2. El multiplicador de Lagrange.
3. Las demandas por factores que resuelven el problemas.
4. Una expresión para la función de costos a partir de las demandas óptimas de cada factor.

Solución.

1. Lo primero será tener en cuenta que para un nivel de producción fijo, digamos Q, tenemos que

$$Q = A \prod_{i=1}^n x_i^{\alpha_i}$$

Luego, el problema de minimización es:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \sum_{i=1}^n w_i x_i \\ \text{s.a} & A \prod_{i=1}^n x_i^{\alpha_i} = Q \end{array}$$

Con esto en mente escribimos la función Lagrangeano

$$L(\boldsymbol{x}, \lambda) = \sum_{i=1}^n w_i x_i - \lambda \left(A \prod_{i=1}^n x_i^{\alpha_i} - Q \right)$$

Para el caso en que la solución es interior ($x_i > 0\ \forall i$) tal que $\lambda = \frac{w_i}{\partial f / \partial x_i}$, las condiciones de primer orden son

$$\frac{\partial L}{\partial x_i} = w_i - \frac{\lambda \alpha_i A \prod_{i=1}^n x_i^{\alpha_i}}{x_i} = 0 \tag{*}$$

$$\frac{\partial L}{\partial \lambda} = Q - A \prod_{i=1}^n x_i^{\alpha_i} = 0 \tag{**}$$

De (*) despejamos x_i

$$x_i = \frac{\lambda \alpha_i A \prod_{i=1}^n x_i^{\alpha_i}}{w_i} = \frac{\lambda \alpha_i Q}{w_i}$$

2. Para obtener el multiplicador definimos $\sum_{i=1}^n \alpha_i = k$ para simplificar la escritura, de la parte anterior tomamos el último resultado, lo reemplazamos en (**) y obtenemos

$$Q^{1-k} = A\lambda^k \cdot \prod_{i=1}^n \left(\frac{\alpha_i}{w_i}\right)^{\alpha_i}$$

Reordenando términos se obtiene lo pedido

$$\begin{aligned}
\lambda^k &= \frac{Q^{1-k}}{A} \cdot \frac{1}{\prod_1^n \left(\frac{\alpha_i}{w_i}\right)^{\alpha_i}} \\
\lambda &= Q^{(1-k)/k}\left(\frac{\prod_{i=1}^n w_i^{\alpha_i/k}}{A^{1/k}\prod_{i=1}^n \alpha_i^{\alpha_i/k}}\right)
\end{aligned}$$

3. Reemplazando el multiplicador en (*) obtenemos la demanda por el insumo j-ésimo.

$$\begin{aligned}
x_j &= Q^{(1-k)/k}\left(\frac{\prod_{i=1}^n w_i^{\alpha_i/\sum_1^n \alpha_i}}{A^{1/k}\prod_{i=1}^n \alpha_i^{\alpha_i/k}}\right) \cdot \frac{\alpha_j Q}{w_j} \\
x_j &= \left(\frac{Q^{1/k}\alpha_i}{w_i}\right) \cdot \frac{\prod_{i=1}^n w_i^{\alpha_i/k}}{A^{1/k}\prod_{i=1}^n \alpha_i^{\alpha_i/k}} \\
x_j &= \left(\frac{Q}{A}\right)^{1/k} \cdot \frac{\prod_{i=1}^n w_i^{\alpha_i/k}}{\prod_{i=1}^n \alpha_i^{\alpha_i/k}} \cdot \frac{\alpha_j}{w_j}
\end{aligned}$$

4. Finalmente para la función de costos multiplicamos la demanda óptima por w_j y agregamos los n términos aplicando la sumatoria $\sum_{j=1}^n(\cdot)$. Se obtiene.

$$C(Q,\boldsymbol{w}) = \sum_{j=1}^n w_j x_j = \left(\frac{Q}{A}\right)^{1/k} \cdot \frac{\prod_{i=1}^n w_i^{\alpha_i/k}}{\prod_{i=1}^n \alpha_i^{\alpha_i/k}} \cdot k$$

$\square$

Ejemplo 7.8. Suponga que la viña Don Pachá se dedica exclusivamente a la producción de vino Carménère, el cual se elabora con n factores productivos y la función de producción es la siguiente:

$$f(\boldsymbol{x}) = A\left(\sum_1^n \alpha_i x_i^\rho\right)^{v/\rho}$$

Donde los parámetros $A,\ \alpha_i,\ \rho,\ v$ son positivos y $\rho \in (0,1)$.

Como en el ejemplo anterior obtendremos:

1. El Lagrangeano del problema.
2. El multiplicador de Lagrange.
3. Las demandas por factores que resuelven el problemas.
4. Una expresión para la función de costos a partir de las demandas óptimas de cada factor.

Solución.

1. La estructura del problema no ha cambiado. Para simplificar el desarrollo algebraico, la función de producción dado un nivel de producción fijo se puede expresar como

$$Q^{\frac{\rho}{v}} = A^{\frac{\rho}{v}} \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right)$$

De esto formamos el Lagrangeano del problema

$$L(\boldsymbol{x}, \lambda) = \sum_{1}^{n} w_i x_i - \lambda \left(A^{\frac{\rho}{v}} \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right) - Q^{\frac{\rho}{v}} \right)$$

Para el caso en que la solución es interior ($x_i > 0 \ \forall i$) tal que $\lambda = \frac{w_i}{\partial f / \partial x_i}$, Las condiciones de primer orden son

$$\frac{\partial L}{\partial x_i} = w_i - \lambda A^{\frac{\rho}{v}} \rho \alpha_i x_i^{\rho - 1} = 0$$

$$\frac{\partial L}{\partial \lambda} = Q^{\frac{\rho}{v}} - A^{\frac{\rho}{v}} \left(\sum_{i=1}^{n} \alpha_i x_i^{\rho} \right) = 0 \qquad (*)$$

De (*) despejamos x_i

$$w_i = \lambda A^{\frac{\rho}{v}} \rho \alpha_i x_i^{\rho - 1}$$

$$x_i = \left(\frac{\lambda A^{\frac{\rho}{v}} \rho a_i}{w_i} \right)^{\frac{1}{1-\rho}} \qquad (**)$$

A partir de este resultado la idea es formar la función de producción nuevamente, para lo cual basta con elevar a ρ, luego multiplicar por α_i, agregar los n términos aplicando la sumatoria $\sum_{i=1}^{n}(\cdot)$ y finalmente multiplicar por $A^{\frac{\rho}{v}}$ para llegar a una expresión equivalente a $Q^{\frac{\rho}{v}}$. Despejando Q se obtiene

$$Q = A^{\frac{1}{1-\rho}} (\lambda \rho)^{\frac{v}{1-\rho}} \left(\sum_{i=1}^{n} a_i^{\frac{1}{1-\rho}} w_i^{\frac{\rho}{\rho - 1}} \right)^{\frac{v}{\rho}}$$

2. De lo anterior, podemos despejar el multiplicador con la finalidad de reemplazar en (**)

$$\lambda^{\frac{1}{1-\rho}} = \frac{Q^{\frac{1}{v}}}{A^{\frac{1}{v(1-\rho)}} \rho^{\frac{1}{1-\rho}} \left(\sum_{i=1}^{n} a_i^{\frac{1}{1-\rho}} w_i^{\frac{\rho}{\rho-1}} \right)^{\frac{1}{\rho}}}$$

3. Reemplazando el multiplicador, que intencionadamente dejamos elevado a $\frac{1}{1-\rho}$ para reemplazar directamente en (**), se obtiene la demanda condicionada por el factor j-ésimo

$$x_j = \left(\frac{Q}{A} \right)^{\frac{1}{v}} \frac{a_j^{\frac{1}{1-\rho}} w_j^{\frac{1}{\rho-1}}}{\left(\sum_{j=1}^{n} a_j^{\frac{1}{1-\rho}} w_j^{\frac{1}{1-\rho}} \right)^{\frac{1}{\rho}}}$$

4. Ahora solo nos falta encontrar la función de costos, para lo cual basta con multiplicar por w_i y finalmente agregar los n términos aplicando la sumatoria $\sum_{j=1}^{n}(\cdot)$. Se obtiene

$$\sum_{j=1}^{n} w_j x_j = \left(\frac{Q}{A}\right)^{\frac{1}{v}} \frac{\displaystyle\sum_{j=1}^{n} a_i^{\frac{1}{1-\rho}} w_i^{\frac{\rho}{\rho-1}}}{\left(\displaystyle\sum_{i=1}^{n} a_i^{\frac{1}{1-\rho}} w_i^{\frac{1}{1-\rho}}\right)^{\frac{1}{\rho}}}$$

En esta última ecuación hay que arreglar las sumatorias que son independientes del índice y se obtiene

$$C(Q,\boldsymbol{w}) = \sum_{i=1}^{n} w_i x_i = \left(\frac{Q}{A}\right)^{\frac{1}{v}} \left(\sum_{i=1}^{n} a_i^{\frac{1}{1-\rho}} w_i^{\frac{\rho}{\rho-1}}\right)^{\frac{\rho-1}{\rho}}$$

□

Ejemplo 7.9. Ahora ilustraremos un caso en que la función objetivo y la restricción son lineales. Esto con la finalidad de dar una interpretación al multiplicador de Lagrange que presentaremos durante el desarrollo del ejercicio y retomaremos más adelante.

Supongamos que los completos de la calle Gorbea son fabricados mediante un proceso que es representable por medio de la siguiente función:

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_i x_i$$

Aunque sea poco realista, ya que da lugar a una perfecta sustitución de insumos, tomaremos esta función para desarrollar un caso en que la solución óptima no es interior y la solución no es directa mediante Lagrangeano.

Como en el ejemplo anterior obtendremos:

1. El Lagrangeano del problema.
2. El multiplicador de Lagrange.
3. Las demandas por factores que resuelven el problemas.
4. Una expresión para la función de costos a partir de las demandas óptimas de cada factor.

Solución.

1. El problema tiene la misma estructura de los dos ejemplos anteriores. Por lo tanto el Lagrangeano del problema es

$$L(\boldsymbol{x},\lambda) = \sum_{i=1}^{n} w_i x_i - \lambda\left(\sum_{i=1}^{n} \alpha_i x_i - Q\right)$$

y las condiciones de primer orden son las siguientes:

$$\frac{\partial L}{\partial x_i} = w_i - \lambda\alpha_i = 0 \qquad \forall i \in \{1,\ldots,n\} \tag{*}$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{n} \alpha_i x_i - Q = 0 \tag{**}$$

Estas condiciones no tienen solución directa.

2. De lo anterior, a partir de (*) tenemos que $\lambda = w_i/\alpha_i$ pero debemos ser criteriosos y considerar

$$\lambda = \min_i \frac{w_i}{\alpha_i}$$

así estamos considerando una única solución del problema cuando existe un único cociente que minimiza λ. Si para algún j se tiene que $w_j/\alpha_j > \lambda$ entonces $x_j = 0$.

3. A partir de (**) tenemos que

$$x_j = \begin{cases} \frac{Q}{\alpha_j} & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En caso de que la solución no sea única cualquier combinación convexa de los distintos x_j es una solución válida del problema. Si D es el conjunto de demandas factibles que resuelven el problema definimos este conjunto de la siguiente forma:

$$D = \left\{ \frac{Q}{\alpha_j} e_j : \frac{w_j}{\alpha_j} \leq \frac{w_i}{\alpha_i} \ \forall i = 1, \ldots, n \right\}$$

donde e_j denota la j-ésima componente de la base canónica de $\mathbb{R}^n$. Así obtenemos que cualquier combinación convexa de demandas factibles da lugar a una demanda factible por factores o insumos. De esta forma la condición para una solución única del problema es que exista sólo un cociente w_i/α_i compatible con

$$\lambda = \min_i \frac{w_i}{\alpha_i}$$

y así garantizamos que la solución no es interior.

Una situación distinta en que no hay solución única es cuando la función objetivo y la restricciíon son iguales, con lo cual cualquier solución interior es óptima y genera el mismo valor en la función objetivo pues no habría un único valor mínimo para el multiplicador.

4. La función de costos está dada por

$$C(Q, \boldsymbol{w}) = Q \cdot \min \left\{ \frac{w_1}{\alpha_1}, \ldots, \frac{w_2}{\alpha_2} \right\}$$

pues habiendo perfecta sustitución de factores y además siendo la solución no interior se elegirá producir utilizando nada más que el factor productivo cuya relación costo-productividad sea menor.

Para fijar ideas es útil pensar en el caso de $\mathbb{R}^2$. Si la pendiente se determina según

$$m_{f(x_1,x_2)} = \frac{\partial f/\partial x_2}{\partial f/\partial x_1}$$

habría que comparar si esta pendiente es mayor, menor o igual a la relación de precios $r_p = p_{x_2}/p_{x_1}$. Se tienen tres casos posibles:

1. $m_{f(x_1,x_2)} > r_p$, entonces la solución óptima es $x_1^* = 0$ y $x_2^* > 0$.
2. $m_{f(x_1,x_2)} < r_p$, entonces la solución óptima es $x_1^* > 0$ y $x_2^* = 0$.
3. $m_{f(x_1,x_2)} = r_p$, entonces la solución óptima es cualquier combinación de valores positivos de x_1 y x_2 que respeten la restricción. Luego, la solución no es única y hay infinitas soluciones para este caso.

$\square$

Ejemplo 7.10. Suponga que los habitantes de Puerto Varas alcanzan cierto nivel de bienestar o utilidad por el consumo de una canasta de n productos. Un estudio de mercado ha revelado que su función de utilidad es la siguiente:

$$u(\boldsymbol{x}) = \prod_{i=1}^{n} x_i^{\alpha_i}$$

Donde $\alpha_i > 0\ \forall i$ y $\sum_{i=1}^{n} \alpha_i = 1$. La restricción al consumo viene dada por un ingreso I que se gasta en su totalidad en productos perfectamente divisibles, los cuales se venden a un precio p_i cada uno.

Obtendremos lo siguiente:

1. El problema de maximización de utilidad del individuo representativo y la función Lagrangeano y las condiciones de primer orden.
2. Una expresión para el multiplicador de Lagrange.
3. La demanda óptima por el producto j-ésimo que resuelve el problema.
4. Una expresión para la función de utilidad en términos de las demandas óptimas.

Solución.

1. El problema a resolver es:

$$\begin{array}{ll} \max\limits_{x} & \prod_{i=1}^{n} x_i^{\alpha_i} \\ \text{s.a} & \sum_{i=1}^{n} p_i x_i = I \end{array}$$

   Del problema de maximización el Lagrangeano corresponde a

$$L(\boldsymbol{x}, \lambda) = \prod_{i=1}^{n} x_i^{\alpha_i} + \lambda \left(I - \sum_{i=1}^{n} p_i x_i \right)$$

   Luego, las condiciones de primer orden respectivas son:

$$\frac{\partial L}{\partial x_i} = \frac{\alpha_i \prod_{i=1}^{n} x_i^{\alpha_i}}{x_i} - \lambda p_i = 0 \qquad (*)$$

$$\frac{\partial L}{\partial \lambda} = I - \sum_{i=1}^{n} p_i x_i = 0 \qquad (**)$$

2. Como $u(\boldsymbol{x}) = \prod_{i=1}^{n} x_i^{\alpha_i}$, vamos a reemplazar en (*) y obtenemos

$$\alpha_i u = \lambda p_i x_i \tag{***}$$

De esto obtenemos λ en términos de $(u, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{x})$

$$\lambda = \frac{\alpha_i u}{p_i x_i}$$

3. Consideremos que $I = \sum_{i=1}^{n} p_i x_i$ y $\sum_{i=1}^{n} \alpha_i = 1$ por lo que en (***) vamos a agregar términos aplicando la sumatoria $\sum_{i=1}^{n}(\cdot)$

$$u = \lambda I$$

Si reemplazamos esto último en (***) obtenemos

$$\alpha_i I = p_i x_i$$

Sólo falta reordenar y obtenemos la demanda por el insumo j-ésimo (los índices son independientes) en términos de $(I, \boldsymbol{\alpha}, \boldsymbol{p})$

$$x_j = \frac{I\alpha_j}{p_j}$$

4. Reemplazamos directamente el último resultado en $u(\boldsymbol{x}) = \prod_{i=1}^{n} x_i^{\alpha_i}$ y obtenemos

$$u(\boldsymbol{x}) = I \prod_{i=1}^{n} \left(\frac{\alpha_i}{p_i}\right)^{\alpha_i}$$

$\square$

Ejemplo 7.11. Suponga que los habitantes de Frutillar alcanzan cierto nivel de bienestar o utilidad por el consumo de una canasta de n productos. Un estudio de mercado ha revelado que su función de utilidad es la siguiente:

$$u(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_i x_i$$

Donde $\alpha_i \geq 0 \ \forall i \in \{1, \dots, n\}$ y $\sum_{i=1}^{n} \alpha_i = 1$. En esta ciudad todos quieren obtener el nivel más alto de utilidad sabiendo que su presupuesto es limitado. Debido a la forma funcional de la utilidad en determinados casos la solución óptima del problema llevará a consumir una cantidad nula de algún producto ante lo cual no se cumple la condición de paralelismo de los gradientes de la función objetivo y las restricciones como se evidencia en cualquier solución interior que podamos encontrar mediante Lagrangeano.

Obtendremos lo siguiente:

1. El problema de maximización de utilidad del individuo representativo y la función Lagrangeano y las condiciones de primer orden.
2. Una expresión para el multiplicador de Lagrange.
3. La demanda óptima por el producto j-ésimo que resuelve el problema.
4. Una expresión para la función de utilidad en términos de las demandas óptimas.

Solución.

1. El problema tiene la misma estructura del ejemplo anterior por lo tanto el Lagrangeano corresponde a
$$L(\boldsymbol{x},\lambda)=\sum_{i=1}^{n}\alpha_i x_i+\lambda\left(I-\sum_{i=1}^{n}p_i x_i\right)$$
y las condiciones de primer orden respectivas son:
$$\frac{\partial L}{\partial x_i}=\alpha_i-\lambda p_i=0 \qquad \forall i\in\{1,\dots,n\} \tag{*}$$
$$\frac{\partial L}{\partial \lambda}=I-\sum_{i=1}^{n}p_i x_i=0 \tag{**}$$
Estas condiciones no tienen solución directa.

2. De lo anterior, a partir de (*) tenemos que $\lambda=\alpha_i/p_i$ pero debemos ser criteriosos y considerar
$$\lambda=\max_i \frac{\alpha_i}{p_i}$$
así estamos considerando una única solución del problema cuando el cociente que maximiza λ es único. Si para algún j se tiene que $\alpha_j/p_j<\lambda$ entonces $x_j=0$.

   A partir de (**) tenemos que
$$x_j=\begin{cases}\frac{I}{p_j} & \text{si } i=j\\ 0 & \text{si } i\neq j\end{cases}$$

3. La solución es única en caso de que no sea interior y debe cumplirse que
$$\frac{\alpha_i}{p_i}\geq\lambda \qquad \forall i\in\{1,\dots,n\}$$
Las soluciones interiores se dan cuando el problema no tiene solución única. En caso de que la solución no sea única cualquier combinación convexa de las demandas es una solución válida del problema. Así si D es el conjunto de demandas factibles que resuelven el problema, definimos D de la siguiente forma:
$$D=\left\{\frac{I}{p_i}e_j : \frac{\alpha_j}{p_j}\geq\frac{\alpha_i}{p_i} \qquad \forall i=1,\dots,n\right\}$$
donde e_j denota la j-ésima componente de la base canónica de $\mathbb{R}^n$. Entonces tendremos que la solución del problema sería cualquier $x_j\in D$.

   De esta forma, la condición para una solución única del problema es que exista solo un cociente α_i/p_i compatible con
$$\lambda=\max_i \frac{\alpha_i}{x_i}$$
y así garantizamos que la solución no es interior.

4. Para obtener una expresión para el nivel de utilidad reemplazamos directamente el último resultado en $u(\boldsymbol{x}) = \sum_{i=1}^{n} \alpha_i x_i$ y obtenemos

$$u(\boldsymbol{x}) = \frac{I\alpha_i}{p_i}$$

en el caso en que la solución no es interior. Si la función objetivo y la restricción son iguales tenemos que

$$u(\boldsymbol{x}) = I \cdot \sum_{i=1}^{1} \gamma_i \frac{\alpha_i}{p_i}$$

donde $\gamma \in [0,1]$ y $\sum_{i=1}^{n} \gamma_i = 1$.

$\square$

Ejemplo 7.12. Suponga que los habitantes de Titirilquén alcanzan cierto nivel de bienestar o utilidad por el consumo de una canasta de n productos. Un estudio de mercado ha revelado que su función de utilidad es la siguiente:

$$u(\boldsymbol{x}, y) = \beta_1 \sum_{i=1}^{n} \alpha_i \ln(x_i) + \beta_2 y$$

Donde $(x_1, \ldots, x_m, y) \in \mathbb{R}^n$, $\beta_i > 0$, $\alpha_i > 0$ $\forall i \in \{1, \ldots, n\}$ y $\sum_{i=1}^{n} \alpha_i = 1$. Las curvas de nivel de esta función, en un caso particular, corresponden al siguiente gráfico

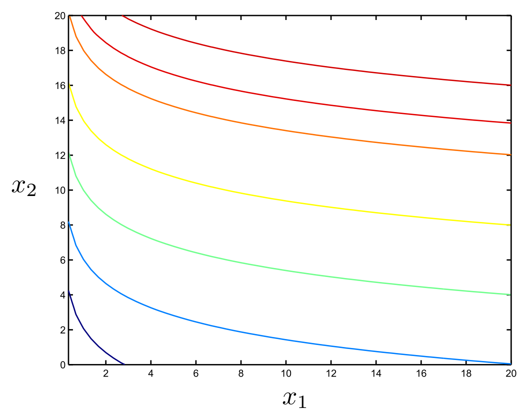


Figura 7.2: Función cuasilineal $f(x_1, x_2) = 0{,}1\ln(x_1) + x_2$

Para fijar ideas, intuitivamente a partir del gráfico del caso particular, es posible deducir que el problema se puede resolver mediante Lagrangeano pero hay que ser cuidadosos puesto que a partir del mismo gráfico se deduce que la solución puede no ser interior.

En esta ciudad todos quieren obtener el nivel más alto de utilidad sabiendo que su presupuesto es limitado. Debido a la forma funcional de la utilidad en determinados casos la solución óptima del problema llevará a consumir una cantidad nula de algún producto ante lo cual no se cumple la condición de paralelismo de los gradientes de la función objetivo y las restricciones como se evidencia en cualquier solución interior que podamos encontrar mediante Lagrangeano.

Ante esta situación:

1. ¿Cómo plantear y resolver el problema de maximización de manera simple?
2. ¿Bajo qué condiciones el método del Lagrangeano nos permite encontrar una solución interior?
3. En caso de existir solución de ambos problemas, ¿se cumple la unicidad de la solución?

Solución.

1. El problema de maximizar utilidad sigue una estructura análoga a la del ejemplo anterior, entonces el Lagrangeano corresponde a

$$L(\boldsymbol{x},\lambda) = \beta_1 \sum_{i=1}^{n} \alpha_i \ln(x_i) + \beta_2 y + \lambda \left(I - \sum_{i=1}^{n} p_i x_i - p_y y \right)$$

Para que exista una solución interior nos basta con tomar las condiciones de primer orden

$$\frac{\partial L}{\partial x_i} = \frac{\beta_1 \alpha_i}{x_i} - \lambda p_i = 0 \ \forall i \in \{1,\dots,n\} \tag{*}$$
$$\frac{\partial L}{\partial y} = \beta_2 - \lambda p_y = 0 \tag{**}$$

De las ecuaciones (*) y (**) se obtienen respectivamente

$$x_i = \frac{\alpha_i \beta_1}{\lambda p_i} \quad \text{y} \quad \lambda = \frac{\beta_2}{p_y}$$

Juntando ambos resultados

$$x_i = \frac{\alpha_i \beta_1 p_y}{\beta_2 p_i}$$

y entonces

$$p_i x_i = \frac{\alpha_i \beta_1 p_y}{\beta_2}$$

Aplicando $\sum_{i=1}^{n}(\cdot)$ a esto último

$$\sum_{i=1}^{n} p_i x_i = \frac{\beta_1 p_y}{\beta_2}$$

Cuando $y > 0$ se tiene que $\sum_{i=1}^{n} p_i x_i \neq I$ y de esto se concluye que

$$y = I - \frac{\beta_1 p_y}{\beta_2}$$

Debemos tener presente que la condición para que esto efectivamente sea una solución interior es $I > \frac{\beta_1 p_y}{\beta_2}$. En caso de que esto último se cumpla con signo de mayor o igual se tendrá que $y \geq 0$.

Si $I < \frac{\beta_1 p_y}{\beta_2}$ se tendrá que en el óptimo $y = 0$ y lo único que nos restaría del problema es resolver la ecuación (*).

Reordenando (*) obtenemos

$$\lambda = \frac{\alpha_i \beta_1}{p_i x_i} \tag{***}$$

Aplicando $\sum_{i=1}^{n}(\cdot)$ obtenemos

$$\lambda I = \beta_1$$

Reemplazando (***) en esto último obtenemos

$$x_i = \frac{\alpha_i I}{p_i}$$

Entonces, la solución del problema es virtud de los parámetros corresponde a

$$x_i = \begin{cases} \dfrac{\alpha_i \beta_1 p_y}{\beta_2 p_i} & \text{si } I \geq \dfrac{\beta_1 p_y}{\beta_2} \\ \\ \dfrac{\alpha_i I}{p_i} & \text{si } I < \dfrac{\beta_1 p_y}{\beta_2} \end{cases} \qquad \text{y} \qquad y = \begin{cases} I - \dfrac{\beta_1 p_y}{\beta_2} & \text{si } \ I \geq \dfrac{\beta_1 p_y}{\beta_2} \\ 0 & \text{si } \ I < \dfrac{\beta_1 p_y}{\beta_2} \end{cases}$$

2. Para una solución interior debe cumplirse que

$$I > \frac{\beta_1 p_y}{\beta_2}$$

Más aún para una solución válida (valores postivos dada la naturaleza del problema) es que la solución la separamos por casos en virtud de los parámetros.

3. La condición encontrada anteriormente garantiza la unicidad de la solución y no hace falta agregar más condiciones.

□

Ejemplo 7.13. Suponga que los habitantes de Salsacia y Conservia alcanzan cierto nivel de bienestar o utilidad por el consumo de una canasta de n productos. El gobierno ha hecho un estudio que revela que la función de utilidad del ciudadano representativo es la siguiente:

$$u(\boldsymbol{x}) = \min_i \{\alpha_i x_i\}$$

Donde $\alpha_i > 0 \ \forall i \in \{1, \ldots, n\}$. Las curvas de nivel de esta función, en un caso particular, corresponden al siguiente gráfico

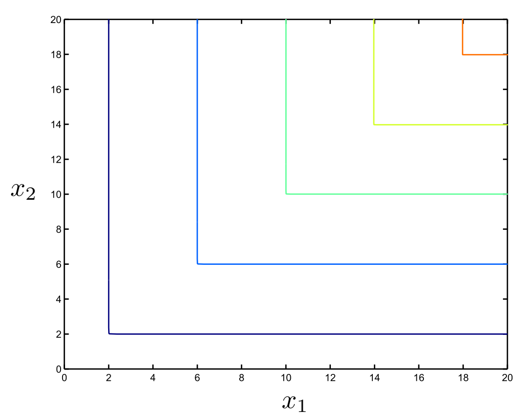


Figura 7.3: Función Leontief $f(x_1, x_2) = \min\{x_1, x_2\}$

Para fijar ideas, intuitivamente a partir del gráfico del caso particular, es posible deducir que el problema no se puede resolver mediante Lagrangeano ya que para $n \geq 2$ la función no es diferenciable.

En esta economía todos quieren obtener el nivel más alto de utilidad sabiendo que su presupuesto es limitado.

Ante esta situación:

1. ¿Cómo plantear y resolver el problema de maximización de manera simple?
2. ¿Bajo qué condiciones el método del Lagrangeano nos permite encontrar una solución interior?
3. En caso de existir solución del problema, ¿se cumple la unicidad de la solución?

Solución.

1. El problema tiene la misma estructura del ejemplo anterior pero no es posible resolver mediante Lagrangeano. Sin embargo, no es muy difícil notar que el vector $\boldsymbol{x}$ de demandas óptimas es tal que
$$\alpha_1 x_1 = \ldots = \alpha_n x_n$$
en general, cada uno de los argumentos es igual a un valor constante $\alpha_i x_i = k$. Luego, se tiene que
$$p_j x_j = \alpha_i x_i \frac{p_j}{\alpha_j}$$
aplicando la sumatoria $\sum_{j=1}^{n}(\cdot)$ se llega a
$$I = \alpha_i x_i \sum_{j=1}^{n} \frac{p_j}{\alpha_j} \;\Rightarrow\; x_i = \frac{I}{\alpha_i \sum_{j=1}^{n} \frac{p_j}{\alpha_j}}$$
2. Para una solución interior bastará con el hecho que $k \neq 0$, que es lo mismo a decir $\alpha_i x_i \neq 0 \; \forall i$. En caso de que $k = 0$ se tendría que $u(\boldsymbol{x}) = 0$ si para algún i se tiene que $\alpha_i x_i = 0$ y de esta forma el vector de demandas óptimas se anularía en todas sus componentes, pero este caso se daría si $I = 0$.
3. En cualquier caso la solución es única.

$\square$

Los casos en que presentamos en los ejemplos 7.9 y 7.11 nos dan la siguiente interpretación del multiplicador de Lagrange: Expresa cuanto baja (resp. sube) el valor de la solución óptima de un problema de minimización (resp. maximización) ante cambios en los parámetros del problema. Esta interpretación es extensible a los demás ejemplos pero la maximizacíon o minimización del cociente que equivale al multiplicador lo deja totalmente en claro. La formalización de esto viene enseguida.

7.2.4. Teorema de la envolvente

Motivación: En diversas áreas de la ingeniería y ciencias nos encontramos con sistemas en los cuales es aplicable un criterio de maximización o minimización. Ya hemos visto formas de resolver esto pero, muchas veces para facilitar el cálculo y ahorrar trabajo cabe preguntarse si existe alguna técnica que nos permita cuantificar cómo cambia la solución óptima ante cambios en los parámetros del sistema o bien, muchas veces tenemos sistemas similares cuya diferencia sólo radica en los parámetros que los definen.

Teorema 7.10. (Teorema de la Envolvente) *Sean* $f : \mathbb{R}^{n+m} \to \mathbb{R}$, $g_i : \mathbb{R}^{n+m} \to \mathbb{R}$ *funciones cóncavas y diferenciables y* $(\boldsymbol{x}, \boldsymbol{c}) \in \mathbb{R}^n \times \mathbb{R}^m$. *Consideremos el problema*

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & f(\boldsymbol{x}, \boldsymbol{c}) \\ s.a & g_i(\boldsymbol{x}, \boldsymbol{c}) = 0 \ \forall i \in \{1, \dots k\} \end{array}$$

y definamos la función valor $V(\boldsymbol{c}) = \max\{f(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) : \boldsymbol{x} \in S\}$.

Como el problema consta únicamente de funciones diferenciables podemos aplicar la función Lagrangeano para encontrar un óptimo. Dado esto el cambio de la función valor, ante cambios en $\boldsymbol{c}$, *equivale al cambio de la función Lagrangeano ante cambios en* $\boldsymbol{c}$. *Es decir,*

$$\frac{\partial V}{\partial c_i}(\boldsymbol{c}) = \frac{\partial L}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{\lambda}(\boldsymbol{c}), \boldsymbol{c})$$

Demostración. Consideremos la función Lagrangeano

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}, \boldsymbol{c}) + \sum_{i=1}^{k} \lambda_i g_i(\boldsymbol{x}, \boldsymbol{c})$$

Sea $\boldsymbol{x}(\boldsymbol{c})$ una solución óptima del problema, tenemos que

$$\frac{\partial L}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{\lambda}(\boldsymbol{c}), \boldsymbol{c}) = \frac{\partial f}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) + \sum_{i=1}^{k} \lambda_i(\boldsymbol{c}) \frac{\partial g_i}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) = 0 \tag{*}$$

Por otra parte,

$$\frac{\partial V}{\partial c_i}(\boldsymbol{c}) = \frac{\partial f}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) \frac{\partial x_i}{\partial c_i}(\boldsymbol{c}) \tag{**}$$

Reemplazando (*) en (**) se obtiene

$$\frac{\partial V}{\partial c_i}(\boldsymbol{c}) = \frac{\partial f}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) - \sum_{i=1}^{k} \lambda_i(\boldsymbol{c}) \left(\sum_{i=1}^{n} \frac{\partial g_i}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) \frac{\partial x_i}{\partial c_i}(\boldsymbol{c}) \right)$$

Pero $g_i(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) = 0 \ \forall i$, entonces

$$\sum_{i=1}^{n} \frac{\partial g_i}{\partial x_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) \frac{\partial x_i}{\partial c_i}(\boldsymbol{c}) + \frac{\partial g_i}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) = 0 \tag{***}$$

Reemplazando (***) en (**) se obtiene

$$\frac{\partial V}{\partial c_i}(\boldsymbol{c}) = \frac{\partial f}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c}) + \sum_{i=1}^{k} \lambda_i(\boldsymbol{c}) \frac{\partial g_i}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{c})$$

que no es otra cosa sino

$$\frac{\partial V}{\partial c_i}(\boldsymbol{c}) = \frac{\partial L}{\partial c_i}(\boldsymbol{x}(\boldsymbol{c}), \boldsymbol{\lambda}(\boldsymbol{c}), \boldsymbol{c})$$

■

Nota 7.5. El caso de minimización es análogo. Basta con tomar f convexa y como $-f$ es cóncava se concluye.

Ejemplo 7.14. Queremos saber cómo cambia la solución óptima ante un cambio en c para el siguiente problema:

$$\begin{array}{ll} \text{máx} & \sqrt{xyz} \\ \text{s.a} & x + 2y + 3z \quad = c \end{array}$$

La función Lagrangeano corresponde a

$$L(x, \lambda) = \sqrt{xyz} - \lambda(x + 2y + 3z - c)$$

Resolviendo las condiciones de primer orden obtenemos

$$\begin{aligned} \frac{\partial L}{\partial x} &= \frac{\sqrt{yz}}{2\sqrt{x}} - \lambda = 0 \\ \frac{\partial L}{\partial y} &= \frac{\sqrt{xz}}{2\sqrt{y}} - 2\lambda = 0 \\ \frac{\partial L}{\partial z} &= \frac{\sqrt{xy}}{2\sqrt{z}} - 3\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= c - x - 2y - 3z = 0 \end{aligned}$$

Reordenando para despejar x, y, z en cada ecuación respectivamente, obtenemos

$$\begin{aligned} x &= \frac{yz}{4\lambda^2} \\ y &= \frac{xz}{16\lambda^2} \\ z &= \frac{xy}{36\lambda^2} \end{aligned}$$

Luego de dividir la primera de estas ecuaciones por la segunda y la segunda con la tercera en forma separada, obtenemos

$$x = 2y = 3z$$

Reemplazamos esto en $\frac{\partial L}{\partial \lambda}$ y obtenemos $c = 3x$. Despejando x, y, z en términos de c obtenemos

$$x(c) = \frac{c}{3} \qquad y = \frac{c}{6} \qquad z = \frac{c}{9}$$

Con esto formamos la función valor y derivamos respecto de c para cuantificar el cambio

$$V(c) = \frac{\sqrt{c^3}}{9\sqrt{2}} \qquad \Rightarrow \qquad \frac{\partial V(c)}{\partial c} = \frac{\sqrt{c}}{6\sqrt{2}}$$

si reemplazamos $x = 2y = 3z$ en $x = \frac{yz}{4\lambda^2}$ obtenemos $c = 72\lambda^2$ y se concluye que $\lambda(c) = \frac{\sqrt{c}}{6\sqrt{2}}$. Con esto último evaluamos

$$\frac{\partial L}{\partial c_i}(x(c), c) = \lambda(c) = \frac{\sqrt{c}}{6\sqrt{2}}$$

Lo que permite concluir que

$$\frac{\partial V(c)}{\partial c} = \frac{\partial L}{\partial c_i}(x(c), c)$$

Dejaremos de *tarea* verificar que el teorema se cumple con los ejemplos de la sección 7.2.3. Para los ejemplos de minimización de costos la forma de proceder es la siguiente:

1. Se define la función valor como la función de costo $C(Q, w)$ y debemos analizar con respecto a Q por ser la constante de la única restricción del los problemas.

2. Se tendrá que $\frac{\partial C(Q,w)}{\partial Q} = \lambda$ exceptuando el ejemplo 7.8, pues en tal caso se debe derivar con respecto a $Q^{\frac{\rho}{v}}$ o de lo contrario llegamos a la conclusión errónea de que el teorema falla. Verifique y explique esto último.

El desarrollo para los ejemplos de maximización de utilidad es análogo.

7.3. Teorema de separación de convexos y lema de Farkas[1]

Los resultados de esta sección nos servirán para demostrar un teorema importante, conocido como Karush-Kuhn-Tucker, que nos permitirá caracterizar las soluciones óptimas cuando tenemos restricciones de igualdad y desigualdad combinadas en un problema.

Teorema 7.11. (Teorema de separación de convexos) *Sean A y B dos conjuntos convexos, disjuntos y diferentes de vacío en $\mathbb{R}^n$. Si A es cerrado y B es compacto existe $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ tal que*

$$\boldsymbol{p} \cdot \boldsymbol{a} < \boldsymbol{p} \cdot \boldsymbol{b} \ \forall (\boldsymbol{a}, \boldsymbol{b}) \in A \times B$$

Demostración. Si A es cerrado definamos la función

$$\begin{aligned} f : B &\to \mathbb{R} \\ \boldsymbol{b} &\mapsto \min_{\boldsymbol{a} \in A} \|\boldsymbol{b} - \boldsymbol{a}\| \end{aligned}$$

la cual nos da la distancia de $\boldsymbol{b} \in B$ a $\boldsymbol{a} \in A$ y además es una función continua. Supongamos que B es compacto, entonces existe $\boldsymbol{b}_0 \in B$ tal que $f(\boldsymbol{b}_0) \leq f(\boldsymbol{b}) \ \forall \boldsymbol{b} \in B$. Sea $\boldsymbol{y} \in A$ tal que $f(\boldsymbol{b}_0) = \|\boldsymbol{b}_0 - \boldsymbol{y}\|$ y como $A \cap B = \varnothing$ son disjuntos entonces el vector

$$\boldsymbol{p} = \frac{\boldsymbol{b}_0 - \boldsymbol{y}}{\|\boldsymbol{b}_0 - \boldsymbol{y}\|}$$

está bien definido y es tal que $\|\boldsymbol{p}\| = 1$.
Como $\boldsymbol{p} \cdot \boldsymbol{p} > 0$ se tiene que

$$\boldsymbol{p} \cdot \frac{\boldsymbol{b}_0 - \boldsymbol{y}}{\|\boldsymbol{b}_0 - \boldsymbol{y}\|} > 0$$

a partir de lo cual se concluye que

$$\boldsymbol{p} \cdot \boldsymbol{y} < \boldsymbol{p} \cdot \boldsymbol{b}_0 \tag{*}$$

En base a esto último debemos demostrar que $\boldsymbol{p} \cdot \boldsymbol{b}_0 \leq \boldsymbol{p} \cdot \boldsymbol{b}$ y $\boldsymbol{p} \cdot \boldsymbol{a} < \boldsymbol{p} \cdot \boldsymbol{y}$.

[1]Tema optativo.

Fijando $\boldsymbol{a} \in A$ definamos la función

$$\begin{aligned} g : [0,1] &\to \mathbb{R} \\ \lambda &\mapsto \|\boldsymbol{b}_0 + \boldsymbol{y} - \lambda(\boldsymbol{a} - \boldsymbol{y})\|^2 \end{aligned}$$

como A es convexo g tiene un mínimo en $\lambda = 0$, entonces

$$\begin{aligned} g(\lambda) &= (\langle \boldsymbol{b}_0 - \boldsymbol{y} - \lambda(\boldsymbol{a} - \boldsymbol{y}), \boldsymbol{b}_0 - \boldsymbol{y} - \lambda(\boldsymbol{a} - \boldsymbol{y})\rangle)^2 \\ &= (\langle \boldsymbol{b}_0 - \boldsymbol{y}, \boldsymbol{b}_0 - \boldsymbol{y}\rangle - 2\lambda\langle \boldsymbol{b}_0 - \boldsymbol{y}, \boldsymbol{a} - \boldsymbol{y}\rangle + \lambda^2\langle \boldsymbol{a} - \boldsymbol{y}, \boldsymbol{a} - \boldsymbol{y}\rangle)^2 \\ \frac{\partial g}{\partial \lambda} &= -2\langle \boldsymbol{b}_0 - \boldsymbol{y}, \boldsymbol{a} - \boldsymbol{y}\rangle + 2\lambda\langle \boldsymbol{a} - \boldsymbol{y}\boldsymbol{a} - \boldsymbol{y}\rangle \end{aligned}$$

luego

$$\left.\frac{\partial g}{\partial \lambda}\right|_{\lambda=0} = (\boldsymbol{b}_0 - \boldsymbol{y}) \cdot (\boldsymbol{y} - \boldsymbol{a}) \geq 0$$

dividiendo por $\|\boldsymbol{b}_0 - \boldsymbol{y}\|$ se concluye que

$$\boldsymbol{p} \cdot \boldsymbol{a} \leq \boldsymbol{p} \cdot \boldsymbol{y} \tag{**}$$

Fijando $\boldsymbol{b} \in B$ definamos la función

$$\begin{aligned} h : &[0,1] \to \mathbb{R} \\ &\lambda \mapsto \|\boldsymbol{b}_0 - \boldsymbol{y} - \lambda(\boldsymbol{b} - \boldsymbol{y})\|^2 \end{aligned}$$

y procediendo de la misma forma con la que se llegó a (**) se concluye que

$$\boldsymbol{p} \cdot \boldsymbol{b}_0 \leq \boldsymbol{p} \cdot \boldsymbol{b} \tag{***}$$

Finalmente (*), (**) y (***) permiten concluir que $\boldsymbol{p} \cdot \boldsymbol{a} < \boldsymbol{p} \cdot \boldsymbol{b}$. ■

Corolario 7.1. *Si A y B son dos conjuntos convexos, disjuntos y diferentes de vacío, entonces el conjunto $C = A \setminus B$ es convexo y no vacío tal que $\mathbf{0} \notin C$. Sobre este resultado pueden darse dos casos:*

1. *$\mathbf{0} \in adh(C)$ y entonces existe $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ tal que $\boldsymbol{p} \cdot \boldsymbol{a} \leq \boldsymbol{p} \cdot \boldsymbol{b}\ \forall(\boldsymbol{a}, \boldsymbol{b}) \in A \times B$.*
2. *$\mathbf{0} \notin adh(C)$ y entonces existe $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ tal que $\boldsymbol{p} \cdot \boldsymbol{a} < \boldsymbol{p} \cdot \boldsymbol{b}\ \forall(\boldsymbol{a}, \boldsymbol{b}) \in A \times B$.*

Demostración.
$\mathbf{0} \in \text{adh}(C)$: Como C es convexo tenemos que el interior de $\text{adh}(C)$ está contenido en C y entonces $\mathbf{0} \notin \text{int}(C)$. Por lo tanto existe una sucesión $\{\boldsymbol{x}_n\}_{n\in\mathbb{N}}$ tal que $\{\boldsymbol{x}_n\} \in \text{adh}(C)$ y $\{\boldsymbol{x}_n\} \to \mathbf{0}$ en la medida que $n \to \infty$. Entonces para cada $n \in \mathbb{N}$ existe un vector $\boldsymbol{p}_n$ tal que $\|\boldsymbol{p}_n\| = 1$ y $\boldsymbol{p}_n \cdot \boldsymbol{a} < \boldsymbol{p}_n \cdot \boldsymbol{b}\ \forall(\boldsymbol{a}, \boldsymbol{b}) \in A \times B$. Como $\{\boldsymbol{p}_n\}_{n\in\mathbb{N}}$ está definida en un compacto tiene al menos una subsucesión convergente, es decir existe $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ tal que $\|\boldsymbol{p}\| = 1$ y $\boldsymbol{p} \cdot \boldsymbol{a} \leq \boldsymbol{p} \cdot \boldsymbol{b}\ \forall(\boldsymbol{a}, \boldsymbol{b}) \in A \times B$.

$\mathbf{0} \notin \text{adh}(C)$: Aplicando el teorema teorema 7.11, se concluye que existe $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ tal que $\boldsymbol{p} \cdot \boldsymbol{c} < 0\ \forall \boldsymbol{c} \in \text{adh}(C)$ y en particular si definimos $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$ con $\boldsymbol{a} \in A$ y $\boldsymbol{b} \in B$ se tiene que $\boldsymbol{p} \cdot \boldsymbol{a} < \boldsymbol{p} \cdot \boldsymbol{b}\ \forall(\boldsymbol{a}, \boldsymbol{b}) \in A \times B$. ■

Nota 7.6. En la última demostración afirmamos que si $\overline{C} = \text{adh}(C)$ entonces $\text{int}(\overline{C}) \subseteq C$. Esta propiedad es intuitivamente clara, sin embargo solo es válida cuando C es convexo. Como contraejemplo tenemos lo siguiente: Sea $C = (0, 1)$ entonces $\text{adh}(C) = [0, 1]$ e $\text{int}(\overline{C}) = (0, 1)$ pero si $C = (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$ entonces $\text{adh}(C) = [0, 1]$ e $\text{int}(\overline{C}) = (0, 1)$ por lo que $\text{int}(\overline{C}) \nsubseteq C$.

Geométricamente el teorema de separación de convexos da la noción de mínima distancia entre dos conjuntos y la existencia de al menos un hiperplano que separa ambos conjuntos. Cuando los conjuntos tienen intersección vacía sabemos que existe un vector $\boldsymbol{p} \neq \boldsymbol{0}$ que separa ambos conjuntos y en $\mathbb{R}^2$, para fijar ideas, resulta útil la siguiente figura:

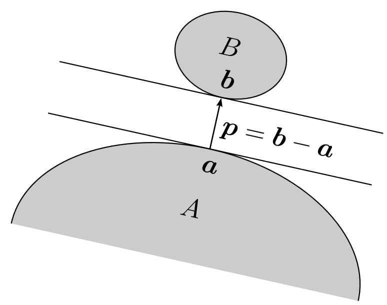


Figura 7.4: Separación estricta.

Siguiendo con el análisis de la separación de convexos, una consecuencia del teorema es que cualquier conjunto cerrado y convexo C puede ser separado de cualquier vector $\boldsymbol{v} \notin C$. Geométricamente tenemos otro caso útil para fijar ideas respecto de un hiperplano separador $H = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{p} \cdot \boldsymbol{x} = \alpha\}$:

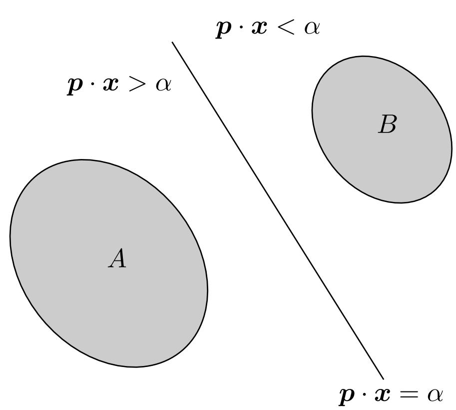


Figura 7.5: Hiperplano separador.

En el caso de conjuntos no convexos estos no necesariamente pueden ser separados. Este caso, el de los no convexos, requiere un análisis aparte.

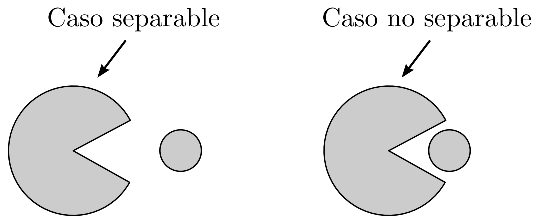


Figura 7.6: Casos no convexos en $\mathbb{R}^2$.

Lema 7.2. (Lema de Farkas) *Sean $\boldsymbol{b} \in \mathbb{R}^n$ y $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. Entonces la desigualdad $\boldsymbol{b}\cdot\boldsymbol{d} \geq 0$ se cumple para todo vector $\boldsymbol{d} \in \mathbb{R}^n$ tal que $A\boldsymbol{d} \geq \boldsymbol{0}$ si y sólo si existe $\boldsymbol{p} \in \mathbb{R}^m_+$ tal que $A^T\boldsymbol{p} = \boldsymbol{b}$.*

Demostración. El enunciado equivale a decir que el sistema $A\boldsymbol{d} \geq \boldsymbol{0},\ \boldsymbol{b}\cdot\boldsymbol{y} < 0$ tiene solución si y sólo si el sistema $A^T\boldsymbol{p} = \boldsymbol{b},\ \boldsymbol{p} \geq \boldsymbol{0}$ no tiene solución.

El sistema $A^T\boldsymbol{p} = \boldsymbol{b},\ \boldsymbol{p} \geq \boldsymbol{0}$ no tiene solución si los conjuntos

$$C_1 = \{\boldsymbol{x} \in \mathbb{R}^n : A^T\boldsymbol{p} = \boldsymbol{x},\ \boldsymbol{p} \geq \boldsymbol{0}\} \quad \text{y} \quad C_2 = \{\boldsymbol{b}\}$$

son disjuntos. Notemos que C_1 y C_2 son cerrados, convexos y diferentes de vacío y además C_2 es compacto. Esto último nos permite aplicar el teorema 7.11 y se concluye que existen $c \in \mathbb{R}^n \setminus \{0\}$ y $\alpha \in \mathbb{R}$ tales que

$$\boldsymbol{c}\cdot\boldsymbol{b} < \alpha\ ,\ \boldsymbol{c}\cdot\boldsymbol{x} > \alpha,\ \forall \boldsymbol{x} \in C_1 \quad \text{y} \quad \boldsymbol{c}\cdot(A^T\boldsymbol{p}) > \alpha,\ \forall \boldsymbol{p} \geq \boldsymbol{0}$$

De esta forma si $\boldsymbol{p} = \boldsymbol{0}$ entonces $\alpha < 0$. Si escogemos $\boldsymbol{p} = (0,\ldots,p_i,\ldots,0)$ con $i = 1,\ldots,m$ tal que $p_i > 0$, entonces $\boldsymbol{c}A^T \geq \boldsymbol{0}$ y en consecuencia $A\boldsymbol{c} \geq \boldsymbol{0}$. Se concluye que $\boldsymbol{d} = \boldsymbol{c}$ es una solución del sistema $A\boldsymbol{d} \geq \boldsymbol{0},\ \boldsymbol{b}\cdot\boldsymbol{d} < 0$.

Si $A^T\boldsymbol{p} = \boldsymbol{b},\ \boldsymbol{p} \geq \boldsymbol{0}$ tiene solución y $A\boldsymbol{y} \geq \boldsymbol{0}$, entonces $\boldsymbol{b}\cdot\boldsymbol{y} = \boldsymbol{p}\cdot(A\boldsymbol{y})$. ■

Ejemplo 7.15. (Precios de no arbitraje) Consideremos un conjunto formado por n activos financieros, cuyos precios al inicio de un periodo de inversión son $p_1,\ldots,p_n$ respectivamente. Al final del periodo de inversión cada activo tendrá un valor $v_1,\ldots,v_n$. Si $x_1,\ldots,x_n$ representa la inversión inicial en cada activo tenemos que:

1. $x_i < 0$ significa que se vendieron x_i unidades del activo i al inicio del periodo.
2. $x_i > 0$ significa que se compraron x_i unidades del activo i al inicio del periodo.

El costo de la inversión inicial es $\boldsymbol{p}\cdot\boldsymbol{x}$ y el valor final de la inversión será de $\boldsymbol{d}\cdot\boldsymbol{x}$ en su totalidad.

Supondremos que el valor final de los activos es incierto al inicio del periodo y que existen m posibles estados de la naturaleza que resultan en distintos valores futuros de la inversión inicial.

En el estado $k \in \{1,\ldots,m\}$ el activo $i \in \{1,\ldots,n\}$ tendrá un valor que se estima en un monto v_i^k por cada unidad adquirida al inicio del periodo. De esta forma, si se adquiere un portafolio $\boldsymbol{x} \in \mathbb{R}^n$ se espera obtener un valor del portafolio igual a $\boldsymbol{d}^k\cdot\boldsymbol{x}$ en el estado k.

La existencia de posibilidades de arbitraje significa que se obtendrá riqueza futura sin comprometer riqueza en el presente. Se dice que existe una posibilidad de arbitraje cuando existe un vector de inversión $\boldsymbol{x}$ con $\boldsymbol{p}\cdot\boldsymbol{x} < 0$ y, en cada estado de la naturaleza, el valor del portafolio es no negativo, es decir, $\boldsymbol{d}^k\cdot\boldsymbol{x} \geq 0$ para cada k.

Definiendo

$$V = \begin{pmatrix} v_1^1 & \dots & v_i^1 \\ & \ddots & \\ v_1^k & \dots & v_i^k \end{pmatrix}$$

Sin que exista la opción de obtener riqueza futura sin comprometer riqueza en el presente tenemos que

$$V\boldsymbol{x} \geq \mathbf{0} \ , \ \boldsymbol{p} \cdot \boldsymbol{x} < 0$$

es infactible o corresponde a un sistema sin solución.

Si el precio de cada activo es es igual al pago de sus pagos futuros, entonces existe $\boldsymbol{y} \in \mathbb{R}^n_{++}$ que cumple $\boldsymbol{p} = V^T\boldsymbol{y}$. Esto último nos dice que no existe posibilidad de arbitraje si

$$V^T\boldsymbol{y} = \boldsymbol{p} \ , \ \boldsymbol{y} \geq 0$$

corresponde a un sistema que tiene solución.

De acuerdo al lema de Farkas esta última posibilidad, en caso de que se cumpla, de inmediato elimina cualquier oportunidad de arbitraje al inicio del periodo.

7.4. Teorema de Karush-Kuhn-Tucker[2]

Motivación: Ahora nos extenderemos a un caso más general de un problema de optimización que se da cuando aparecen restricciones de igualdades y desigualdades en forma combinada. El teorema de los multiplicadores de Lagrange exige independencia lineal entre los gradientes de la función objetivo y las restricciones en el óptimo lo cual resulta en una condición demasiado fuerte y restrictiva. Ahora presentaremos un teorema que no requiere dicha hipótesis y que no tiene el inconveniente que surge al no poder trabajar con un número de restricciones mayor a la dimensión del espacio debido a la condición de independencia lineal.

Daremos dos demostraciones del teorema de Karush-Kuhn-Tucker: Mediante el teorema de la función implícita como ya se hizo para el teorema de los multiplicadores de Lagrange y mediante el Lema de Farkas la cual nos da una interpretación mucho más clara del signo de los multiplicadores.

7.4.1. Condiciones de primer orden para extremos restringidos

Definición 7.10. (Espacios normal y tangente)
Por analogía con la definición 7.8 definiremos la superficie

$$S = \{g_i(\boldsymbol{x}) = 0, \, h_j(\boldsymbol{x}) = 0 \ \forall i \in I, \ j \in J(\boldsymbol{x}_0)\}$$

en torno a $\boldsymbol{x}_0$ y así tenemos que los espacios normal y tangente corresponden a

$$N_S(\boldsymbol{x}_0) = \langle \{\nabla g_i(\boldsymbol{x}_0), \, \nabla h_j(\boldsymbol{x}_0)\} \rangle \ \forall i \in I, \ j \in J$$

$$T_S(\boldsymbol{x}_0) = N_S(\boldsymbol{x}_0)^{\perp} = \{v \in \mathbb{R}^n : v \cdot \nabla g_i(\boldsymbol{x}_0) = 0 \ \wedge \ v \cdot \nabla h_j(\boldsymbol{x}_0) \leq 0 \ \forall i \in I, \ j \in J\}$$

En caso de que alguna restricción del problema (7.2) sea de la forma $h_j(\boldsymbol{x}_0) < 0$ para algún j, se tiene que esta no participa en la estructura de S. Este hecho motiva la siguiente definición:

[2] Tema optativo.

Definición 7.11. (Espacio linealizante)
El espacio linealizante de S en $\boldsymbol{x}_0$ corresponde al conjunto

$$L_S(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \; \forall i \in I, \; \nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \; \forall j \in J(\boldsymbol{x}_0)\}$$

donde $J(\boldsymbol{x}_0) = \{j \in J : h_j(\boldsymbol{x}_0) = 0\}$ corresponde a las restricciones de menor o igual que efectivamente participan en la estructura de S.

El siguiente lema es un primer acercamiento a la demostración del teorema:

Lema 7.3. (Análogo a lema 7.1)
Consideremos el siguiente problema de optimización

$$\begin{array}{llll} P') & \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) & \\ & s.a & g_i(\boldsymbol{x}) = 0 & \forall i \in I \\ & & h_j(\boldsymbol{x}) = 0 & \forall j \in J(\boldsymbol{x}_0) \end{array} \tag{7.4}$$

Este problema se obtiene de (7.2) cuando consideramos sólo restricciones que se cumplen con igualdad en $\boldsymbol{x}_0$. Entonces, $\forall v \in T_S(\boldsymbol{x}_0)$ existe $\sigma : A \subseteq \mathbb{R} \to \mathbb{R}^n$, $0 \in A$ con $g_i(\sigma(t)) = 0, h_j(\sigma(t)) = 0$ $\forall t \in A$, $\sigma(t) \in S$ $\forall t \in A$ y además $\sigma(0) = \boldsymbol{x}_0$ tal que $\sigma'(0) = \boldsymbol{d}$. Es decir,

$$T_S(\boldsymbol{x}_0) = \{\sigma'(0) : \sigma(t) \in S \; \forall t, \sigma(0) = \boldsymbol{x}_0\}$$

Demostración. Definamos $q = k + p$. Sean $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$ y $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{n-q-1}\}$ una base ortonormal del espacio

$$\langle \{\boldsymbol{d}, \nabla g_i(\boldsymbol{x}_0), \nabla h_j(\boldsymbol{x}_0)\} \rangle^{\perp} \; \forall i \in I, \; j \in J(\boldsymbol{x}_0)$$

y consideremos el sistema de $n-1$ ecuaciones

$$\begin{array}{rcl} g_1(\boldsymbol{x}) & = & 0 \\ & \vdots & \\ h_p(\boldsymbol{x}) & = & 0 \\ (\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{b}_1 & = & 0 \\ & \vdots & \\ (\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{b}_{n-q-1} & = & 0 \end{array}$$

El vector $\boldsymbol{x}_0$ satisface las ecuaciones, y la matriz jacobiana del sistema está dada por

$$\begin{bmatrix} \nabla g_1(\boldsymbol{x}_0) \\ \vdots \\ \nabla h_p(\boldsymbol{x}_0) \\ \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_{n-q-1} \end{bmatrix}$$

que es una matriz de rango $n-1$ por lo que podemos seleccionar $n-1$ columnas tales que la matriz resultante sea invertible y gracias al teorema de la función implícita (teorema 5.3) podemos despejar $n-1$ variables en función de la restante en una vecindad de $\boldsymbol{x}_0$. Entonces existe $\sigma : A \to \mathbb{R}^n$, con

$A = (-\varepsilon, \varepsilon)$, tal que $g_i(\sigma(t)) = 0,\ h_j(\sigma(t)) = 0\ \forall t \in A$ y $(\sigma(t) - \boldsymbol{x}_0) \cdot \boldsymbol{b}_h = 0\ \forall t,\ h \in \{1, \ldots, n-q-1\}$. Derivando las ecuaciones anteriores con respecto a t y evaluando en $t = 0$ se obtiene que

$$\begin{aligned}
\sigma'(0) \cdot \boldsymbol{b}_h &= 0 \quad \forall h \in \{1, \ldots, n-q-1\} \\
\nabla g_i(\boldsymbol{x}_0) \cdot \sigma'(0) &= 0 \quad \forall i \in I \\
\nabla h_j(\boldsymbol{x}_0) \cdot \sigma'(0) &= 0 \quad \forall j \in J(x_0)
\end{aligned}$$

lo cual implica que $\sigma'(0) \parallel \boldsymbol{d}$. Definiendo $\bar{\sigma}(t) = \sigma(\|\boldsymbol{d}\| t / \|\sigma'(0)\|)$ tenemos que $\bar{\sigma}'(0) = \boldsymbol{d}$, lo que nos da el lema. ■

Teorema 7.12. (Teorema de Karush-Kuhn-Tucker, primera versión)
Una condición necesaria para que un vector $\boldsymbol{x}_0 \in S$ sea solución del problema 7.2 tal que $\boldsymbol{x}_0$ es regular es que exista un vector $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^k \times \mathbb{R}^m_+$ tal que

$$\nabla_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla f(\boldsymbol{x}_0) + \sum_{i \in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j \in J} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0 \tag{7.5}$$

$$\mu_j h_j(\boldsymbol{x}_0) = 0\ \forall j \in J \tag{7.6}$$

$$g_i(\boldsymbol{x}_0) = 0\ \forall i \in I \tag{7.7}$$

$$h_j(\boldsymbol{x}_0) \leq 0\ \forall j \in J \tag{7.8}$$

$$\mu_j \geq 0\ \forall j \in J \tag{7.9}$$

Cuando $f,\ g_i,\ h_i$ son convexas las condiciones descritas son suficientes para que $\boldsymbol{x}_0$ sea un mínimo de f en S.

Demostración. (Utilizando el teorema de la función implícita)
La condición (7.5) se demuestra mediante el lema 7.3. Consideremos la superficie S de la definición 7.10 y así $\forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0)$ existe $\sigma : A \subseteq \mathbb{R} \to \mathbb{R}^n$, $0 \in A$ con $g_i(\sigma(t)) = 0,\ h_j(\sigma(t)) \leq 0\ \forall (t \in A)$ y además $\sigma(0) = \boldsymbol{x}_0$ tal que $\sigma'(0) = \boldsymbol{d}$.

Si $\boldsymbol{x}_0$ es solución del problema 7.2 entonces existe $r > 0$ tal que $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})\ \forall \boldsymbol{x} \in B(\boldsymbol{x}_0, r) \cap S$. Juntando esto con lo anterior tenemos que

$$f(\boldsymbol{x}_0) = f(\sigma(0)) \leq f(\sigma(t))\ \forall t \in (-\varepsilon_1, \varepsilon_1)$$

Dado esto, $t_0 = 0$ minimiza $f(\sigma(t))$ entonces

$$\left.\frac{d}{dt} f(\sigma(t))\right|_{t=0} = \nabla f(\boldsymbol{x}_0) \cdot \sigma'(0) = 0$$

y dado que $\boldsymbol{x}_0$ es un vector regular de 7.2 también lo es de S por como definimos $J(\boldsymbol{x}_0)$.

En consecuencia, todo vector $\boldsymbol{d}$ en el espacio tangente $T_S(\boldsymbol{x}_0)$ a S en $\boldsymbol{x}_0$ está dado por $\boldsymbol{d} = \sigma'(0)$ para alguna función $\sigma'(t)$ y de esta forma $\nabla f(\boldsymbol{x}_0)$ es una combinación lineal de $\{\nabla g_i(\boldsymbol{x}_0),\ \nabla h_j(\boldsymbol{x}_0)\}\ \forall i \in I,\ j \in J(\boldsymbol{x}_0)$.

El lema 7.3 nos dice que si $\boldsymbol{x}_0$ es un mínimo local de f restringido a S, lo anterior es equivalente a que $\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0$ para $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$, es decir $\nabla f(\boldsymbol{x}_0) \in N_S(\boldsymbol{x}_0)$ lo que implica que

$$\nabla f(\boldsymbol{x}_0) + \sum_{i \in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j \in J(\boldsymbol{x}_0)} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0$$

De esto último tenemos que $\nabla f(\boldsymbol{x}_0) \in T_S(\boldsymbol{x}_0)^\perp$ ya que por definición de $T_S(\boldsymbol{x}_0)$ y $N_S(\boldsymbol{x}_0)$ tenemos que $T_S(\boldsymbol{x}_0) = N_S(\boldsymbol{x}_0)^\perp$ lo que nos lleva a $T_S(\boldsymbol{x}_0)^\perp = N_S(\boldsymbol{x}_0)$ y así $\mu_j = 0\ \forall j \notin J(\boldsymbol{x}_0)$, entonces

$$\nabla f(\boldsymbol{x}_0) + \sum_{i \in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j \in J} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0$$

La condición (7.6) se demuestra directamente a partir de lo siguiente: Dado cualquier j tal que $1 \leq j \leq m$ pueden ocurrir uno de los siguientes casos:

1. $j \in J(\boldsymbol{x}_0)$ y en tal caso $h_j(\boldsymbol{x}_0) = 0$
2. $j \notin J(\boldsymbol{x}_0)$ y en tal caso $\mu_j = 0$

Finalmente, la condición (7.9) se demuestra por contradicción. Supongamos que $\mu_j < 0$ y $h_j(\boldsymbol{x}_0) < 0$ para algún j tal que $1 \leq j \leq m$ y $j \notin J(\boldsymbol{x}_0)$, como el multiplicador no se anula debería cumplirse que $j \in J$ ya que $j \notin J(\boldsymbol{x}_0)$ en caso de que $\mu_j = 0$. De acuerdo al lema 7.3 existe $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$ tal que $\boldsymbol{d} \cdot \nabla h_j(\boldsymbol{x}_0) < 0$ y así

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} = \sum_{i \in I} \lambda_i (\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d}) + \sum_{j \in J} \mu_j (\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d})$$

Se obtiene

$$\begin{aligned}
\left.\frac{d}{dt} f(\sigma(t))\right|_{t=0} &= \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \\
&= -\sum_{i \in I} \lambda_i (\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d}) - \sum_{j \in J} \mu_j (\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d}) \\
&= -\mu_j (\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d}) \\
&< 0
\end{aligned}$$

y se contradice el hecho de que $t_0 = 0$ minimiza $\sigma(t)$ y que $\boldsymbol{x}_0$ es solución de 7.2 por lo que existiría $t \in (-\varepsilon_2, \varepsilon_2)$ tal que $\sigma(t) \leq \sigma(0)$. ■

Bajo ciertas condiciones se tiene que $T_S(\boldsymbol{x}_0) = L_S(\boldsymbol{x}_0)$, una de estas condiciones es la siguiente:

Definición 7.12. (Condiciones de Mangasarian-Fromovitz)
Diremos que $\boldsymbol{x}_0 \in S$ es regular si se cumplen

1. $\{\nabla g_i(\boldsymbol{x}_0) : i \in I\}$ es linealmente independiente.
2. $\exists \boldsymbol{d} \in \mathbb{R}^n$ tal que $\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I$ y además $\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} < 0 \ \forall j \in J(\boldsymbol{x}_0)$.

Para lo cual una condición suficiente (y no necesaria) es que

$$\{\nabla g_i(\boldsymbol{x}_0)\}_{i \in I} \cup \{\nabla h_j(\boldsymbol{x}_0)\}_{j \in J(\boldsymbol{x}_0)}$$

sea linealmente independiente.

Antes de presentar la segunda versión del teorema nos será de enorme utilidad el siguiente resultado:

Teorema 7.13. *Bajo condiciones de regularidad de Mangasarian-Fromovitz se tiene que* $T_S(\boldsymbol{x}_0) = L_S(\boldsymbol{x}_0)$

Demostración.
$T_S(\boldsymbol{x}_0) \subset L_S(\boldsymbol{x}_0)$: Si $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$ entonces existen $\boldsymbol{x}_n \in S$ y $t_n \to 0$ tales que $\boldsymbol{d} = \lim_{n \to \infty} (\boldsymbol{x}_n - \boldsymbol{x}_0)/t_n$. Luego para todo $j \in J(\boldsymbol{x}_0)$ se cumple que $h_j(\boldsymbol{x}_k) - h_j(\boldsymbol{x}_0) \leq 0$ y la expansión de Taylor sobre h_j da el siguiente resultado:

$$h_j(\boldsymbol{x}_0) + \nabla h_j(\boldsymbol{x}_0) \cdot (\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\| R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \leq 0$$

Diviendo por t_n y tomando límite cuando $n \to \infty$ se tendrá que $R_1 \to 0$ y como $h_j(\boldsymbol{x}_0) = 0$ para el caso $j \in J(\boldsymbol{x}_0)$ se concluye que

$$\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} \leq 0$$

Es análogo para verificar que $\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0$.

$L_S(\boldsymbol{x}_0) \subset T_S(\boldsymbol{x}_0)$: Sea $\boldsymbol{d} \in L_S(\boldsymbol{x}_0)$ y consideremos el sistema no lineal en las variables $(t, \boldsymbol{u})$

$$g_i(\boldsymbol{x} + t\boldsymbol{d} + A\boldsymbol{u}) = 0, \ i \in I$$

donde A es la matriz cuyas columnas son $\nabla g_i(\boldsymbol{x}_0)$. El punto $(t, \boldsymbol{u}) = \boldsymbol{0}$ es solución del sistema y la matriz Jacobiana respecto de $\boldsymbol{u}$ en $\boldsymbol{0}$ es $A^T A$ la cual es invertible de acuerdo a las condiciones de regularidad establecidas. Por medio del teorema de la función implícita es posible concluir que existe una solución $\boldsymbol{u}(t)$ diferenciable en torno a $t = 0$ con $\boldsymbol{u}(t) = \boldsymbol{0}$. De esta forma se tiene que

$$\boldsymbol{x}_0(t) = \boldsymbol{x}_0 + t\boldsymbol{d} + A\boldsymbol{u}(t)$$

satisface la igualdad $g_i(\boldsymbol{x}_0(t)) = 0 \ \forall i \in I, \ t \ B(\boldsymbol{0}, r)$. De esto es posible concluir que

$$\frac{d}{dt} g_i(\boldsymbol{x}_0(t))(\boldsymbol{0}) = \nabla g_i(\boldsymbol{x}_0) \cdot \left(\boldsymbol{d} + A \frac{d\boldsymbol{u}(t)}{dt} \right) \bigg|_{t=0}$$

a partir de $\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0$ y la invertibilidad $A^T A$ se concluye que

$$\left. \frac{d\boldsymbol{u}(t)}{dt} \right|_{t=0} = \boldsymbol{0} \text{ y por lo tanto } \left. \frac{d\boldsymbol{x}(t)}{dt} \right|_{t=0} = \boldsymbol{d}$$

La trayectoria $\boldsymbol{x}_0(t)$ satisface las igualdades $g_i(\boldsymbol{x}_0(t)) = 0$.

Para las desigualdades tenemos dos alternativas:

$\boldsymbol{d} \in L_S(\boldsymbol{x}_0)$ suponiendo desigualdades estrictas: Tenemos que $\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} < 0 \ \forall j \in J(\boldsymbol{x}_0)$. En tal caso, dado que

$$h_j(\boldsymbol{x}_0(t)) = h_j(\boldsymbol{x}_0) + t\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} + tr(t)$$

con $r(t) \to 0$ en la medida que $t \to 0$, se tiene que $h_j(\boldsymbol{x}_0(t)) < 0 \ \forall j \in J$ con $t \approx 0$.
Lo anterior demuestra que $\boldsymbol{x}_0(t) \in S$ cuando $t \approx 0$ y dado que $\boldsymbol{d} = \lim_{n\to\infty}(\boldsymbol{x}_n - \boldsymbol{x}_0)/t_n$ para cualquier sucesión $\{t_n\}_{n\in\mathbb{N}} \to 0^+$ y se concluye que $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$.

$\boldsymbol{d} \in L_S(\boldsymbol{x}_0)$ sin suponer desigualdades estrictas: Para cada $\varepsilon > 0$, el vector $\boldsymbol{d}_\varepsilon = \boldsymbol{d} + \varepsilon \boldsymbol{d}_0 \ \in L_S(x_0)$ y satisface las desigualdades estrictas $\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d}_\varepsilon < 0 \ \forall j \in J(\boldsymbol{x}_0)$. De la segunda parte de la demostración se concluye que $\boldsymbol{d}_\varepsilon \in T_S(\boldsymbol{x}_0)$. Como $T_S(\boldsymbol{x}_0)$ es cerrado, en la medida que $\varepsilon \to 0$ se concluye que $\boldsymbol{d} \in T_S(\boldsymbol{x}_0)$. ■

Teniendo el teorema anterior en mente podemos enunciar la segunda versión del teorema y la demostración se simplifica considerablemente.

Teorema 7.14. (Teorema de Karush-Kuhn-Tucker, segunda versión)
Una condición necesaria para que $\boldsymbol{x}_0 \in S$ sea solución del problema 7.2 tal que $T_S(\boldsymbol{x}_0) = L_S(\boldsymbol{x}_0)$ (por ejemplo bajo condiciones de Mangasarian-Fromovitz) es que exista un vector $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^k \times \mathbb{R}^m_+$

tal que

$$\nabla_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla f(\boldsymbol{x}_0) + \sum_{i\in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j\in J} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0 \tag{7.10}$$

$$\mu_j h_j(\boldsymbol{x}_0) = 0 \ \forall j \in J \tag{7.11}$$

$$g_i(\boldsymbol{x}_0) = 0 \ \forall i \in I \tag{7.12}$$

$$h_j(\boldsymbol{x}_0) \leq 0 \ \forall j \in J \tag{7.13}$$

$$\mu_j \geq 0 \ \forall j \in J \tag{7.14}$$

Cuando f, g_i, h_i son convexas las condiciones descritas son suficientes para que $\boldsymbol{x}_0$ sea un mínimo de f en S.

Demostración. (Utilizando el lema de Farkas)
De la condición necesaria se tiene que $\boldsymbol{x}_0$ es un mínimo local de f en S si

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0 \ \forall \boldsymbol{d} \in T_S(\boldsymbol{x}_0)$$

Luego, cumplíendose las condiciones de Mangasarian-Fromovitz se tiene que $T_S(\boldsymbol{x}_0) = L_S(\boldsymbol{x}_0)$ y será posible aplicar el lema de Farkas (lema 7.2).

El lema nos dice que dado $A\boldsymbol{d} \geq 0$ se cumple que $\boldsymbol{b} \cdot \boldsymbol{d} \geq 0$. La condición $\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I$ se puede expresar convenientemente como

$$\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0 \quad \text{y} \quad -\nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0 \ \forall i \in I$$

y dado que $-\nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0, \ \forall j \in J(\boldsymbol{x}_0)$ podemos definir la matriz

$$A = \begin{bmatrix} \nabla g_i(\boldsymbol{x}_0)^T \\ -\nabla g_i(\boldsymbol{x}_0)^T \\ -\nabla h_j(\boldsymbol{x}_0)^T \end{bmatrix}_{\substack{i\in I \\ j\in J(\boldsymbol{x}_0)}} \quad \text{y el vector } \boldsymbol{b} = \nabla f(\boldsymbol{x}_0)$$

También el lema nos dice que lo anterior es válido si dado $\boldsymbol{b} \in \mathbb{R}^n$ existe $\boldsymbol{p} \in \mathbb{R}^m_+$ tal que

$$A^T \boldsymbol{p} \geq \boldsymbol{b}, \ \boldsymbol{p} \geq 0$$

De acuerdo al teorema 7.11 el vector $\boldsymbol{p} \neq \boldsymbol{0}$ puede ser escogido con componentes $(p_i^1, p_i^2, p_j)_{i\in I, \ j\in J(\boldsymbol{x}_0)}$ donde no todas las componentes son nulas, entonces

$$\nabla f(\boldsymbol{x}_0) = -\sum_{i\in I}(p_i^2 - p_i^1)\nabla g_i(\boldsymbol{x}_0) - \sum_{j\in J(\boldsymbol{x}_0)} p_j \nabla h_j(\boldsymbol{x}_0)$$

Definiendo $\lambda_i = p_i^2 - p_i^1$ y $\mu_j = p_j$ se tiene

$$\nabla f(\boldsymbol{x}_0) = -\sum_{i\in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) - \sum_{j\in J(\boldsymbol{x}_0)} \mu_j \nabla h_j(\boldsymbol{x}_0)$$

En caso de que $\nabla f(\boldsymbol{x}_0) = 0$ no se cumple la condición suficiente (y no necesaria) de independencia lineal de los gradientes de las restricciones, este caso lleva a

$$\sum_{i\in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j\in J(\boldsymbol{x}_0)} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0 \ \Rightarrow \ \sum_{j\in J(\boldsymbol{x}_0)} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0$$

Para el caso $\nabla f(\boldsymbol{x}_0) \neq 0$ se obtiene

$$\nabla f(\boldsymbol{x}_0) + \sum_{i \in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j \in J(\boldsymbol{x}_0)} \mu_j \nabla h_j(\boldsymbol{x}_0) = 0 \tag{*}$$

Finalmente, si alguna restricción h_j no participa en la estructura de S entonces $h_j(\boldsymbol{x}_0) < 0 \ \forall j \notin J(\boldsymbol{x}_0)$ y se puede definir $\mu_j = 0 \ \forall j \notin J(\boldsymbol{x}_0)$. Entonces existe $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^k \times \mathbb{R}^m_+$ tal que

$$\begin{aligned} \nabla f(\boldsymbol{x}_0) + \sum_{i \in I} \lambda_i \nabla g_i(\boldsymbol{x}_0) + \sum_{j \in J} \mu_j \nabla h_j(\boldsymbol{x}_0) &= 0 \\ \mu_j h_j(\boldsymbol{x}_0) &= 0 \ \forall j \in J \\ \mu_j &\geq 0 \ \forall j \in J \end{aligned}$$

■

En las últimas condiciones encontradas la primera se conoce como condición de primer orden y la segunda como condición de holgura complementaria y exige que los multiplicadores asociados a las restricciones inactivas sean nulos. En otras palabras, sólo las restricciones activas deben ser tomadas en consideración.

Definición 7.13. El conjunto de todos los multiplicadores $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^k \times \mathbb{R}^m_+$ que satisfacen las condiciones del teorema de Karush-Kuhn-Tucker será denotado $\Lambda(\boldsymbol{x}_0)$. Nótese que bajo la condición de independencia lineal el conjunto $\Lambda(\boldsymbol{x}_0)$ se reduce a un único vector.

Con lo ya discutido en la sección 7.2 el teorema 7.12 también es válido para la maximización si consideramos restricciones de la forma $h_j(\boldsymbol{x}_0) \geq 0 \ \forall j \in J$. Una condición necesaria para que un vector regular y factible $\boldsymbol{x}_0$ sea máximo de f en S es que se cumplan las condiciones que describe el teorema y cuando $f, \ g_i, \ h_j$ son cóncavas la condición es suficiente. Los multiplicadores μ_j siguen siendo no negativos para el caso de maximización.

Respecto de los signos de los multiplicadores, no hay restricción de signo para los multiplicadores asociados a las restricciones de igualdad. En el caso de restricciones de desigualdad, el signo de los multiplicadores depende de si estamos empleando un criterio de maximización o minimización y si las restricciones se dejan como mayores o iguales a cero.
Si el problema es de minimización sujeto a restricciones de menor o igual entonces los multiplicadores son mayores o iguales a cero y se incluyen con signo positivo en la función Lagrangeano. Si estamos maximizando con restricciones de mayor o igual no cambia el signo de los multiplicadores.

7.4.2. Condiciones de segundo orden para extremos restringidos

Definición 7.14. De manera similar a como se hizo en la sección 7.2.2 definiremos el conjunto de direcciones críticas para un problema de minimización como

$$K(\boldsymbol{x}_0) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I, \ \nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} \leq 0 \ \forall j \in J(\boldsymbol{x}_0), \ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0\}$$

Teorema 7.15. (Condición necesaria de segundo orden)
Sea $\boldsymbol{x}_0$ un mínimo local del problema 7.2 bajo las condiciones de Mangasarian-Fromovitz. Entonces, para todo $\boldsymbol{d} \in K(\boldsymbol{x}_0)$ existe un multiplicador $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \Lambda(\boldsymbol{x}_0)$ tal que

$$\boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))\boldsymbol{d} \geq 0$$

Demostración. Por simplicidad se demostrará considerando hipótesis de independencia lineal. Sea $\boldsymbol{d} \in K(\boldsymbol{x}_0)$, definamos $J_d(\boldsymbol{x}_0) = \{j \in J(\boldsymbol{x}_0) : \nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0\}$, y consideremos el sistema

$$\begin{aligned} g_i(\boldsymbol{x}) &= 0 \ \forall i \in I \\ h_j(\boldsymbol{x}) &= 0 \ \forall j \in J_d(\boldsymbol{x}_0) \end{aligned}$$

De manera análoga a la demostración del teorema 7.13, en virtud del teorema de la función implćita, es posible encontrar una trayectoria $\boldsymbol{x}(t)$ de clase $\mathcal{C}^2$ en torno a $t = 0$, la cual cumple que $g_i(\boldsymbol{x}(t)) = 0 \ \forall i \in I \ h_j(\boldsymbol{x}(t)) = 0 \ \forall i \in J_d(\boldsymbol{x}_0)$, con $\boldsymbol{x}(0) = 0$ y

$$\frac{d}{dt}\boldsymbol{x}(0) = \boldsymbol{d}$$

Ya que cada una de las restricciones $h_j(\boldsymbol{x}_0) \leq 0$ son continuas para $t \approx 0$ tendremos que $h_j(\boldsymbol{x}(t)) < 0 \ \forall j \in J \setminus J_d(\boldsymbol{x}_0)$, de manera que $\boldsymbol{x}(t) \in S$ para todo $t \to 0$.

Por otra parte, como $\boldsymbol{x}_0$ es mínimo local del problema 7.2 y $\boldsymbol{x}(t)$ es factible, entonces $t = 0$ genera un mínimo local de f en S si evaluamos f en $\boldsymbol{x}(t)$. Luego, dado que $\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0$ ya que d es arbitrario, se deduce que

$$0 \leq \frac{d^2}{dt^2} f(\boldsymbol{x}(0)) = \boldsymbol{d}^T (H_x f(\boldsymbol{x}_0))\boldsymbol{d} + \nabla f(\boldsymbol{x}_0) \cdot \frac{d^2}{dt^2}\boldsymbol{x}(0) \tag{*}$$

Análogamente obtenemos

$$\begin{aligned} 0 \leq \frac{d^2}{dt^2} g_i(\boldsymbol{x}(0)) &= \boldsymbol{d}^T (H_x g_i(\boldsymbol{x}_0))\boldsymbol{d} + \nabla g_i(\boldsymbol{x}_0) \cdot \frac{d^2}{dt^2}\boldsymbol{x}(0) \ \forall i \in I \\ 0 \leq \frac{d^2}{dt^2} h_j(\boldsymbol{x}(0)) &= \boldsymbol{d}^T (H_x h_j(\boldsymbol{x}_0))\boldsymbol{d} + \nabla h_j(\boldsymbol{x}_0) \cdot \frac{d^2}{dt^2}\boldsymbol{x}(0) \ \forall j \in J_d(\boldsymbol{x}_0) \end{aligned}$$

multiplicando estas igualdades por λ_i y μ_j respectivamente y sumando a (*) obtenemos

$$0 \leq \boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))\boldsymbol{d}$$

■

Teorema 7.16. (Condición sufiente de segundo orden)
Sea $\boldsymbol{x}_0$ un vector factible del problema 7.2 y supongamos que cumple que para todo $\boldsymbol{d} \in K(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$ existe $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \Lambda(\boldsymbol{x}_0)$ tal que

$$\boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))\boldsymbol{d} > 0$$

Entonces $\boldsymbol{x}_0$ es un mínimo local estricto.

Demostración. Procederemos por contradicción. Si $\boldsymbol{x}_0$ no es un mínimo local estricto entonces existe $\boldsymbol{x}_n \in S$ tal que $\boldsymbol{x}_n \to \boldsymbol{x}_0$ y $f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0)$ para $\boldsymbol{x}_n \neq \boldsymbol{x}_0$.

Definamos

$$\boldsymbol{d}_n = \frac{\boldsymbol{x}_n - \boldsymbol{x}_0}{\|\boldsymbol{x}_n - \boldsymbol{x}_0\|}$$

tal que $\|\boldsymbol{d}_n\| = 1$ y entonces $\{\boldsymbol{d}_n\}$ es acotada por lo que tiene una subsucesión convergente. En efecto, podemos suponer que converge a $\boldsymbol{d}_0 \in K(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$. Probemos entonces que $\boldsymbol{d}_0$ es una dirección crítica. De manera similar a la demostración del teorema 7.2 el suponer que $\boldsymbol{d}_0 \in K(\boldsymbol{x}_0) \setminus \{0\}$ conduce a que $\boldsymbol{d}_0 \in T_S(\boldsymbol{x}_0) \setminus \{\boldsymbol{0}\}$ y además se tiene que $T_S(\boldsymbol{x}_0) \subset L_S(\boldsymbol{x}_0)$. Por otra parte, la expansión de Taylor de $f(\boldsymbol{x}_n)$ está dada por

$$f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\| R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \leq 0$$

esto más la condición $f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0)$ conducen a

$$\nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\| R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \leq 0$$

diviendo por $\|\boldsymbol{x}_n - \boldsymbol{x}_0\|$ y tomando límite cuando $n \to \infty$ resulta

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d}_0 \leq 0$$

Sea $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ el multiplicador asociado a $\boldsymbol{d}$ en la condición suficiente de segundo orden. Como $L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ es de clase $\mathcal{C}^2$ en $\boldsymbol{x}$ y $(\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \Lambda(\boldsymbol{x}_0)$, entonces la expansión de Taylor de $L(\boldsymbol{x}_n, \boldsymbol{\lambda}, \boldsymbol{\mu})$ corresponde a

$$L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}) + \frac{1}{2}(\boldsymbol{x}_n - \boldsymbol{x}_0)^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))(\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2 R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \tag{*}$$

Por otra parte, de las condiciones de Karush-Kuhn-Tucker se tiene que

$$L(\boldsymbol{x}_n, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f(\boldsymbol{x}_n) \leq f(\boldsymbol{x}_0) = L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu})$$

lo cual determina que $L(\boldsymbol{x}_n, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu})$ y este resultado más la ecuación (*) conducen a

$$\frac{1}{2}(\boldsymbol{x}_n - \boldsymbol{x}_0)^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))(\boldsymbol{x}_n - \boldsymbol{x}_0) + \|\boldsymbol{x}_n - \boldsymbol{x}_0\|^2 R_1(\|\boldsymbol{x}_n - \boldsymbol{x}_0\|) \leq 0$$

diviendo por $\|\boldsymbol{x}_n - \boldsymbol{x}_0\|$ y tomando límite cuando $n \to \infty$ resulta

$$\boldsymbol{d}^T (H_x L(\boldsymbol{x}_0, \boldsymbol{\lambda}, \boldsymbol{\mu}))\boldsymbol{d} \leq 0$$

lo cual es una contradicción. ■

De la misma forma que como sucede en el teorema 7.8, para el teorema anterior basta con que el Hessiano del Lagrangreano solo con respecto a x sea definido positivo en el conjunto de direcciones críticas y no en cualquier dirección arbitraria. Las condiciones suficientes de segundo orden no requiren condiciones de Mangasarian-Fromovitz ni tampoco de independencia lineal.

Para el caso de maximización basta con cambiar el sentido de las desigualdades para la condición necesaria y la condición suficiente de segundo orden pues en este caso el conjunto de direcciones críticas se define

$$K(\boldsymbol{x}_0) = \{\boldsymbol{d} \in \mathbb{R}^n : \nabla g_i(\boldsymbol{x}_0) \cdot \boldsymbol{d} = 0 \ \forall i \in I, \ \nabla h_j(\boldsymbol{x}_0) \cdot \boldsymbol{d} \leq 0 \ \forall j \in J(\boldsymbol{x}_0), \ \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{d} \geq 0\}$$

7.4.3. Interpretación económica del teorema de Karush-Kuhn-Tucker

En el caso de las empresas competitivas resulta razonable suponer que su objetivo es la maximización de beneficios económicos. Simplificando la realidad de una empresa podemos decir que esta es monoproductora y elige un nivel de producción $y \in \mathbb{R}_+$, dado el precio de venta $p \in \mathbb{R}_{++}$ de su producto el cual es determinado exógenamente.

Podemos decir que la empresa tiene una función de producción (o de transformación de insumos) $y = f(\boldsymbol{x})$ donde $y > 0$ es el nivel de producción escogido y $\boldsymbol{x} \in \mathbb{R}^n_+$ es el vector de insumos utilizados en la producción. Esta función la definiremos como

$$f(\boldsymbol{x}) = \text{máx}(y) \text{ tal que } (y, \boldsymbol{x}) \in Y, \ Y = \{(y, \boldsymbol{x}) \in \mathbb{R}_+ \times \mathbb{R}^n_+ \ f(\boldsymbol{x}) \geq y\}$$

Cada insumo tiene un costo unitario $w_i > 0$, lo que se traduce en un vector de costos $\boldsymbol{w} \in \mathbb{R}^n_+$. Entonces la empresa debe resolver las siguientes ecuaciones:

$$\begin{aligned}
\pi(p,\boldsymbol{w}) &= \sup_{\boldsymbol{x}\in\mathbb{R}^n_+} pf(\boldsymbol{x}) - \boldsymbol{w}\cdot\boldsymbol{x} \\
\boldsymbol{x}(p,\boldsymbol{w}) &= \arg\max_{\boldsymbol{x}\in\mathbb{R}^n_+}\ pf(\boldsymbol{x}) - \boldsymbol{w}\cdot\boldsymbol{x}
\end{aligned}$$

$\pi(p,\boldsymbol{w})$ corresponde a la función de beneficios y da cuenta de la diferencia entre ingreso y costos. No hemos hecho los suficientes supuestos para asegurar que se alcanza un máximo beneficio (por ejemplo: $y(p) \neq \varnothing$), entonces no podemos reemplazar sup por máx ignorando que son diferentes conceptualmente hablando. En particular podría darse el caso en que $\pi(p) \to +\infty$, lo cual sucedería si Y no es acotado.

Lo que sigue es sobre la base de dos supuestos fuertes: Existe una única decisión de producción que efectivamente maximiza beneficios y las posibilidades de producción están acotadas .

La maximización de beneficios se puede separar en dos partes:

1. Primero se encuentra una combinación de factores que permita producir a costos mínimo dado un nivel de producción y.

$$\begin{aligned}
\text{Función de costo: } C(y,\boldsymbol{w}) &= \inf_{\boldsymbol{x}:f(\boldsymbol{x})\geq y} \boldsymbol{w}\cdot\boldsymbol{x} \\
\text{Demanda por factores: } X(y,\boldsymbol{w}) &= \arg\min_{\boldsymbol{x}:f(\boldsymbol{x})\geq y} \boldsymbol{w}\cdot\boldsymbol{x} \\
&= \{\boldsymbol{x} : f(\boldsymbol{x}) \geq y \ \wedge\ \boldsymbol{w}\cdot\boldsymbol{x} = C(y,\boldsymbol{w})\}
\end{aligned}$$

2. Encontrar un nivel de producción que maximice la diferencia entre los ingresos y los costos

$$\max_{y\geq 0}\ py - C(y,\boldsymbol{w})$$

De lo anterior ahora podemos pasar a minimización de costos. Supongamos que Y sólo considera como restricción que se puede producir a lo más $\overline{y}$ dada la capacidad tecnológica de la empresa:

$$Y = \{(q,\boldsymbol{x}) : \boldsymbol{x}\in\mathbb{R}^n_+ \ \wedge\ f(\boldsymbol{x}) \geq q\}$$

El problema de minimización de costos es simila al de maximización de beneficios sobre el conjunto Y si tomamos $y \in Y$ y además

$$\begin{aligned}
\pi(p,\boldsymbol{w}) &= py - C(y,\boldsymbol{w}) \\
y(\boldsymbol{w}) &= (y, X(y,\boldsymbol{w}))
\end{aligned}$$

Una construcción adecuada del problema de maximización de beneficios es la siguiente:

$$\begin{aligned}
\max_{\boldsymbol{x}\in\mathbb{R}^n_+} &\quad pf(\boldsymbol{x}) - \boldsymbol{w}\cdot\boldsymbol{x} \\
\text{s.a} &\quad x_i \geq 0\ \forall i \in \{1,\dots,n\}
\end{aligned}$$

por obvio que parezca la restricción es importante porque valores negativos de $\boldsymbol{x}$ carecen de sentido en este contexto. Luego construimos el Lagrangeano del problema

$$L(\boldsymbol{x},\boldsymbol{w},p,\boldsymbol{\mu}) = pf(\boldsymbol{x}) - \boldsymbol{w}\cdot\boldsymbol{x} + \boldsymbol{\mu}\cdot\boldsymbol{x}$$

De lo cual se obtienen las siguientes condiciones

Condición de primer orden:	$p\nabla f(\boldsymbol{x}_0) - w + \mu = 0$	
Holgura complementaria:	$\mu_i x_0^i = 0$	$\forall i \in \{1, \ldots, n\}$
No negatividad:	$\mu_i \geq 0$	$\forall i \in \{1, \ldots, n\}$
Restricción inicial:	$x_0^i \geq 0$	$\forall i \in \{1, \ldots, n\}$

Nota 7.7. De acuerdo a la notación empleada $\boldsymbol{x}_0 = (x_0^1, \ldots, x_0^n) \in \mathbb{R}^n$, es decir x_0^i es la i-ésima coordenada de $\boldsymbol{x}_0$.

La consecuencia de estas condiciones consideradas en forma simultánea es la siguiente: Para todo i se cumplirá que

$$p\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} \leq w_i$$

y además si $x_0^i > 0$ sabemos que $\mu_i = 0$ entonces para cualquier insumo que se utilice en cantidades positivas se tendrá

$$p\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} = w_i$$

Esta condición en los libros de economía se expresa: "El valor de producto marginal del factor en competencia perfecta es lo que está dispuesta a pagar la firma (o empresa) por dicho factor". Si el costo del factor i es superior a $p\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i}$ dicho factor no se utilizará.

Centrémonos ahora en el problema de minimización de costos. Una construcción adecuada del problema es la siguiente:

$$\begin{array}{cl} \displaystyle\min_{\boldsymbol{x} \in \mathbb{R}^n_+} & \boldsymbol{w} \cdot \boldsymbol{x} \\ \text{s.a} & f(\boldsymbol{x}) \geq y \\ & x_i \geq 0 \ \forall i \in \{1, \ldots, n\} \end{array}$$

por obvio que parezca la restricción es importante pues valores negativos de $\boldsymbol{x}$ carecen de sentido en este contexto. Luego construimos el Lagrangeano del problema (asumiremos $f(\boldsymbol{x}) = y$)

$$L(\boldsymbol{x}, \boldsymbol{w}, y, \lambda, \boldsymbol{\mu}) = \boldsymbol{w} \cdot \boldsymbol{x} + \lambda(q - f(\boldsymbol{x})) - \boldsymbol{\mu} \cdot \boldsymbol{x}$$

Imponiendo las mismas condiciones que en el caso anterior llegamos a lo siguiente:

Condición de primer orden:	$\boldsymbol{w} - \lambda\nabla f(\boldsymbol{x}_0) - \boldsymbol{\mu} = 0$	
Holgura complementaria:	$\mu_i x_0^i = 0$	$\forall i \in \{1, \ldots, n\}$
No negatividad:	$\mu_i \geq 0$	$\forall i \in \{1, \ldots, n\}$
Restricción inicial:	$f(\boldsymbol{x}_0) = y\ ,\ x_0^i \geq 0$	$\forall i \in \{1, \ldots, n\}$

Estas condiciones en forma simultánea conducen a

$$\lambda\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} \leq w_i \ \forall i \in \{1, \ldots, n\}$$

y para todo $x_0^i > 0$ se tiene

$$\lambda\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} = w_i$$

en este caso el multiplicador de Lagrange da cuenta de que el uso de los insumos o factores guarda relación con la proporción que existe entre el costo del factor y su productividad marginal.

Centrémonos ahora en el problema de obtener una cantidad óptima de producción (es similar a maximización de beneficios). Una construcción adecuada del problema es la siguiente:

$$\begin{array}{ll} \max\limits_{y\in\mathbb{R}_+} & py - C(y,\boldsymbol{w}) \\ \text{s.a} & y \geq 0 \end{array}$$

Luego construimos el Lagrangeano del problema

$$L(y,p,\boldsymbol{w},\mu) = py - C(y,\boldsymbol{w}) + \mu y$$

Imponiendo las mismas condiciones que en el caso anterior llegamos a lo siguiente:

$$\begin{array}{ll} \text{Condición de primer orden:} & p - \frac{\partial C(y_0,\boldsymbol{w})}{\partial y} + \mu = 0 \\ \text{Holgura complementaria:} & \mu y_0 = 0 \\ \text{No negatividad:} & \mu \geq 0 \\ \text{Restricción inicial:} & y_0 \geq 0 \end{array}$$

Estas condiciones en forma simultánea conducen a

$$p \leq \frac{\partial C(y_0,\boldsymbol{w})}{\partial y}$$

si $q_0 > 0$ se tiene

$$p = \frac{\partial C(y_0,\boldsymbol{w})}{\partial y}$$

lo cual nos dice que la situación óptima se logra con un nivel de producción tal que el costo marginal de producir una unidad adicional es igual al precio de venta del producto.

Notemos que si la solución óptima de los tres problemas es $(y_0,\boldsymbol{x}_0) > 0$ entonces

$$p\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} = w_i \;\wedge\; \lambda\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} = w_i \;\Rightarrow\; p\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} = \lambda\frac{\partial f(\boldsymbol{x}_0)}{\partial x_i} \;\Rightarrow\; p = \lambda$$

y como $p = \frac{\partial C(y_0,\boldsymbol{w})}{\partial y}$ obtenemos

$$p = \lambda = \frac{\partial C(y_0,\boldsymbol{w})}{\partial y}$$

Lo cual tiene una conclusión muy importante: El multiplicador de Lagrange corresponde al precio de equilibrio en el óptimo y la oferta de la firma, bajo los supuestos de los cuales partimos, se determina a partir de su costo marginal cuando se iguala con el precio de venta. Sin embargo esto último corresponde a una condición necesaria y no suficiente para determinar la oferta de la firma. Debemos tener presente que las condiciones de KKT son sólo necesarias.

7.4.4. Ejemplos

Ejemplo 7.16. Francisca tiene una función de utilidad (o nivel de bienestar) por el consumo de frutillas (x) y cerezas (y) de la forma

$$u(x,y) = x^{1/2} + \frac{y}{4}$$

Lo que le interesa es maximizar utilidad. Sabemos que el precio de venta de cada fruta es de 1 unidad monetaria y que cada día cuenta con un presupuesto de una unidad monetaria exclusivamente para estas frutas.

Resuelva el problema que enfrenta Francisca utilizando multiplicadores de Lagrange y compare su resultado con la solución que se encuentra mediante Kuhn-Tucker.

Solución. El problema es el siguiente:

$$\begin{array}{ll} \max\limits_{x,y} & x^{1/2} + \dfrac{y}{4} \\ \text{s.a} & x + y = 1 \end{array}$$

Entonces el Lagrangeano nos queda de la siguiente forma:

$$L(x, y, \lambda) = x^{1/2} + \frac{y}{4} + \lambda(1 - x - y)$$

Las condiciones de primer orden son las siguientes

$$\begin{array}{lcl} \dfrac{\partial L}{\partial x} = 0 & \Leftrightarrow & \dfrac{1}{2x_0^{1/2}} - \lambda = 0 \\ \dfrac{\partial L}{\partial y} = 0 & \Leftrightarrow & \dfrac{1}{4} - \lambda = 0 \\ \dfrac{\partial L}{\partial \lambda} = 0 & \Leftrightarrow & 1 - x_0 - y_0 = 0 \end{array}$$

Si igualamos λ en las primeras dos ecuaciones se obtiene $x_0 = 4$ y reemplazando en la tercera ecuación se obtiene $y_0 = -3$. Claramente esta solución carece de sentido dada la naturaleza del problema.

Si utilizamos Kuhn-Tucker podemos incorporar restricciones de no negatividad al problema y nos queda de la siguiente manera:

$$\begin{array}{ll} \max\limits_{x,y} & x^{1/2} + \dfrac{y}{4} \\ \text{s.a} & 1 - x - y \geq 0 \\ & x \geq 0 \\ & y \geq 0 \end{array}$$

Entonces el Lagrangeano nos queda de la siguiente forma:

$$L(x, y, \mu_1, \mu_2, \mu_3) = x^{1/2} + \frac{y}{4} + \mu_1(1 - x - y) + \mu_2 x + \mu_3 y$$

Las condiciones de primer orden son las siguientes

$$\begin{array}{lcl} \dfrac{\partial L}{\partial x} = 0 & \Leftrightarrow & \dfrac{1}{2x_0^{1/2}} - \mu_1 + \mu_2 = 0 \\ \dfrac{\partial L}{\partial y} = 0 & \Leftrightarrow & \dfrac{1}{4} - \mu_1 + \mu_3 = 0 \end{array}$$

y además el teorema nos dice que debemos imponer las siguientes condiciones

$$\begin{array}{rcl} \mu_1(1 - x - y) & = & 0 \\ \mu_2 x & = & 0 \\ \mu_3 y & = & 0 \end{array}$$

Supongamos que $\mu_1 > 0$ entonces de acuerdo al teorema se tiene que la primera restricción se cumple con igualdad.

Sin levantar nuestro supuesto digamos que $\mu_2 = 0$ y $\mu_3 > 0$ entonces $x_0 > 0$ e $y_0 = 0$ y en este caso tenemos que tras reemplazar en la restricción la única alternativa es $x_0 = 1$ dado que $y_0 = 0$. El otro caso de interés es decir que $\mu_1 > 0$ y $\mu_3 = 0$ entonces $x_0 = 0$ e $y_0 > 0$ lo cual lleva a que la única alternativa es $y_0 = 1$ dado que $x_0 = 0$. Resulta que el par $(1, 0)$ del primer caso genera un valor en la función objetivo igual a $f(1, 0) = 1$ mientras que el par $(0, 1)$ del segundo caso genera un valor en la función objetivo igual a $f(0, 1) = 1/4$ por lo que nos quedamos con la primera solución. Otro caso es suponer que $\mu_2, \mu_3 > 0$ y en tal caso la utilidad es cero. □

Ejemplo 7.17. Suponga que Karina es la nueva gerente de la empresa Control de Gestión S.A. La primera medida que ha implementado como política de la empresa es maximizar el ingreso por ventas teniendo en cuenta que los beneficios económicos (ingreso menos costos) de la empresa no pueden bajar de un nivel fijo m.

El costo de publicidad en revistas corresponde a un valor $a \in \mathbb{R}_+$. Siendo $I(y, a) = py + f(a)$ el ingreso que recibe la empresa, p el precio de venta, $y \in \mathbb{R}_+$ el nivel de producción y f una función creciente en el nivel de publicidad el cual es $a \in \mathbb{R}_+$.

Sea $C(y)$ el costo de producción asociado a fabricar una cantidad y determinada. Se sabe que las funciones de ingreso y costo son funciones $\mathcal{C}^1$ y ambas son crecientes en el nivel de producción, es decir, $\partial C/\partial y > 0$ y $\partial I/\partial a > 0$.

Un estudio de mercado encargado por la gerente reveló que $y^* > 0$. ¿Cómo resolvería un problema que permita encontrar un nivel de producción y un gasto en publicidad (y^*, a^*) óptimo?

Solución. El problema de la empresa nos queda de la siguiente forma:

$$\begin{array}{ll} \max\limits_{y,a} & I(y,a) \\ \text{s.a} & \pi(y,a) = I(y,a) - C(y) - a \geq m \\ & a \geq 0 \\ & y \geq 0 \end{array}$$

Asumamos que existe un par (y^*, a^*) con $y^* > 0$ correspondiente a una solución óptima. El lagrangeano corresponde a:

$$L(y, a, \mu_1, \mu_2, \mu_3) = I(y, a) + \mu_1(I(y, a) - C(y) - a - m) + \mu_2 a + \mu_3 y$$

Por el teorema de KKT sabemos que $\mu_2 y^* = 0$ e $y^* > 0$ implican $\mu_2 = 0$ por lo que sin pérdida de generalidad el Lagrangeano del problema se puede escribir:

$$L(y, a, \mu_1, \mu_2) = I(y, a) + \mu_1(I(y, a) - C(y) - a - m) + \mu_2 a$$

Si la solución óptima cumple la condición de IL o MF, las condiciones de primer orden generan las siguientes ecuaciones

$$\frac{\partial L}{\partial y} = (1 + \mu_1)\frac{\partial I}{\partial y} - \mu_1 \frac{\partial C}{\partial y} = 0 \tag{*}$$

$$\frac{\partial L}{\partial a} = (1 + \mu_1)\frac{\partial I}{\partial a} - \mu_1 + \mu_2 = 0 \tag{**}$$

y además el teorema nos dice que debemos imponer las siguientes condiciones

$$\begin{aligned} \mu_1(m + I(y,a) - C(y) - a) &= 0 \\ \mu_2 a &= 0 \end{aligned} \quad (***)$$

Como $\partial I/\partial a > 0$ y $\mu_1 \geq 0$ en la ecuación (**) tenemos que $\mu_2 - \mu_1 < 0$. Como $\mu_2 \geq 0$, μ_1 debe ser estrictamente positivo. Por lo tanto, de (***) se deduce que $\pi(y^*, a^*)$ por lo que el beneficio percibido es el mínimo aceptable.

Como $\mu_1 > 0$ y $\partial C/\partial y > 0$ de la ecuación (*) se deduce que $\partial I/\partial y > 0$ lo que nos dice que el ingreso marginal el positivo en el óptimo.

Por otra parte, tenemos que el beneficio marginal $\partial\pi/\partial y$ en el óptimo es negativo pues de otra forma el nivel de producción necesariamente tendría que ser menor y los beneficios serían mayores pues nos encontramos en la situación de mínimo beneficio aceptable, la ecuación (*) verifica esto pues

$$\begin{aligned} (1+\mu_1)\frac{\partial\pi}{\partial y} &= (1+\mu_1)\left(\frac{\partial I}{\partial y} - \frac{\partial C}{\partial y}\right) \\ &= \frac{\partial L(y^*, a^*)}{\partial y} - \frac{\partial C}{\partial y} \\ &= 0 - \frac{\partial C}{\partial y} \\ &< 0 \end{aligned}$$

En consecuencia, verificamos que el nivel de producción y^* es mayor que el que se escogería en una situación de maximización de beneficios. □

Ejemplo 7.18. Suponga que Rodrigo asume como el nuevo gerente de Energía Renovable S.A. la cual instala calefactores solares modelo Básico (x) y modelo Premium (y). Lo que necesita es determinar su plan de producción óptimo. Según un estudio el beneficio por cada unidad de producto instalada está dado por

Básico	$800 - x - y$
Premium	$2000 - x - 3y$

Donde x e y son las cantidades totales que instala de cada producto. Para instalación se requiere mano de obra y uso de maquinaria según la siguiente tabla

	Recurso (horas/unidad)	
Producto	**Mano de obra**	**Maquinaria**
Básico	8	7
Premium	3	6
Disponibilidad (horas/mes)	1200	2100

Se pide:

1. Plantear el problema de optimización y formular el Lagrangeano del problema.

2. Determinar todas las soluciones óptimas y/o factibles.

Solución. El problema es

$$
\begin{array}{ll}
\max\limits_{x,y} & (800-x-y)x+(2000-x-3y)y \\
\text{s.a} & 1200-8x-3y \geq 0 \\
 & 2100-7x-6y \geq 0 \\
 & x \geq 0 \\
 & y \geq 0
\end{array}
$$

Antes de resolver debemos determinar si las condiciones de KKT son al menos necesarias para la maximización. Para esto tenemos que el hessiano de la función corresponde a

$$
H = \begin{pmatrix} -2 & -2 \\ -2 & -6 \end{pmatrix}
$$

que es definido negativo porque $(-1)^1 \cdot |H_1| > 0$ y $(-1)^2 \cdot |H_2| > 0$ y así la función es cóncava por lo que las condiciones de KKT son suficientes.

El Lagrangeano corresponde a

$$
\begin{aligned}
L(x,\boldsymbol{\mu}) &= (800-x-y)x+(2000-x-3y)y \\
&\quad +\mu_1(1200-8x-3y)+\mu_2(2100-7x-6y)+\mu_3(x-0)+\mu_4(y-0)
\end{aligned}
$$

Entonces las condiciones de primer orden generan el siguiente sistema

$$
\begin{cases}
800-2x-2y-8\mu_1-7\mu_2+\mu_3 &= 0 \\
2000-2x-6y-3\mu_1-6\mu_2+\mu_4 &= 0
\end{cases}
$$

y además el teorema nos dice que debemos imponer las siguientes condiciones

$$
\begin{aligned}
\mu_1(1200-8x-3y) = 0 \wedge \mu_1 \geq 0 \\
\mu_2(2100-7x-6y) = 0 \wedge \mu_2 \geq 0 \\
\mu_3 x = 0 \wedge \mu_3 \geq 0 \\
\mu_4 y = 0 \wedge \mu_4 \geq 0
\end{aligned}
$$

Sea $x_0 > 0$ e $y_0 > 0$, dejaremos de *tarea* los casos $(x_0 = 0, y_0 > 0)$ y $(x_0 > 0, y_0 = 0)$. Tenemos que $\mu_3 = 0$ y $\mu_4 = 0$ y sobre este resultado pueden pasar cuatro cosas

Caso 1: $(1200-8x-3y=0)$ y $(2100-7x-6y=0)$
Entonces $\mu_1 > 0$ y $\mu_2 > 0$ por lo que las condiciones de KKT se reducen a

$$
\begin{aligned}
800-2x-2y-8\mu_1-7\mu_2 = 0 \\
2000-2x-6y-3\mu_1-6\mu_2 = 0 \\
1200-8x-3y = 0 \\
2100-7x-6y = 0 \\
x > 0 \\
y > 0
\end{aligned}
$$

Para resolver se desarrolla el sistema

$$
\begin{pmatrix} 2 & 2 & 8 & 7 \\ 2 & 6 & 3 & 6 \\ 8 & 3 & 0 & 0 \\ 7 & 6 & 0 & 0 \end{pmatrix}
\begin{pmatrix} x \\ y \\ \mu_1 \\ \mu_2 \end{pmatrix}
=
\begin{pmatrix} 800 \\ 2000 \\ 1200 \\ 2100 \end{pmatrix}
$$

y se obtiene

$$x_0 = 33,\bar{3} \qquad y_0 = 311,\bar{1} \qquad \mu_1 = 7,407 \qquad \mu_2 = 7,407$$

que es una solución óptima y factible.

Caso 2: $(1200 - 8x - 3y \neq 0)$ y $(2100 - 7x - 6y = 0)$
Entonces $\mu_1 = 0$ y $\mu_2 > 0$ por lo que las condiciones de KKT se reducen a

$$\begin{aligned} 800 - 2x - 2y - 7\mu_2 &= 0 \\ 2000 - 2x - 6y - 6\mu_2 &= 0 \\ 1200 - 8x - 3y &> 0 \\ 2100 - 7x - 6y &= 0 \\ x &> 0 \\ y &> 0 \end{aligned}$$

Para resolver se desarrolla el sistema

$$\begin{pmatrix} 2 & 2 & 7 \\ 2 & 6 & 6 \\ 8 & 3 & 0 \\ 7 & 6 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ \mu_2 \end{pmatrix} = \begin{pmatrix} 800 \\ 2000 \\ 2100 \end{pmatrix}$$

y se obtiene

$$x_0 = 39,394 \qquad y_0 = 304,04 \qquad \mu_2 = 16,162$$

que es una solución óptima pero no factible.

Caso 3: $(1200 - 8x - 3y = 0)$ y $(2100 - 7x - 6y \neq 0)$
Entonces $\mu_1 > 0$ y $\mu_2 = 0$ por lo que las condiciones de KKT se reducen a

$$\begin{aligned} 800 - 2x - 2y - 8\mu_1 &= 0 \\ 2000 - 2x - 6y - 3\mu_1 &= 0 \\ 1200 - 8x - 3y &= 0 \\ 2100 - 7x - 6y &> 0 \\ x &> 0 \\ y &> 0 \end{aligned}$$

Para resolver se desarrolla el sistema

$$\begin{pmatrix} 2 & 2 & 8 \\ 2 & 6 & 3 \\ 8 & 3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ \mu_1 \end{pmatrix} = \begin{pmatrix} 800 \\ 2000 \\ 1200 \end{pmatrix}$$

y se obtiene

$$x_0 = 31,373 \qquad y_0 = 316,34 \qquad \mu_1 = 13,072$$

que es una solución óptima pero no factible.

Caso 4: $(1200 - 8x - 3y \neq 0)$ y $(2100 - 7x - 6y \neq 0)$

Entonces $\mu_1 = 0$ y $\mu_2 = 0$ por lo que las condiciones de KKT se reducen a

$$\begin{aligned} 800 - 2x - 2y &= 0 \\ 2000 - 2x - 6y &= 0 \\ 1200 - 8x - 3y &> 0 \\ 2100 - 7x - 6y &> 0 \\ x &> 0 \\ y &> 0 \end{aligned}$$

Para resolver se desarrolla el sistema

$$\begin{pmatrix} 2 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 800 \\ 2000 \end{pmatrix}$$

y se obtiene

$$x_0 = 100 \qquad y_0 = 300$$

que es una solución óptima pero no factible. □

Ejemplo 7.19. Aldo tiene la posibilidad de invertir en n activos $x_1, \dots, x_n$ que ofrecen una tasa de retorno aleatoria $r_1, \dots, r_n$ respectivamente y cada activo tiene una tasa de retorno promedio $r_0^i = E(r_i)$ para $i = 1, \dots, n$ y la covarianza del activo i con el activo j es σ_{ij} para $j = 1, \dots, n$. El portafolio $y = \sum_{i=1}^n r_i x_i$ formado por todos los activos tiene una tasa media de retorno dada por

$$E(y) = \sum_{i=1}^n r_0^i x_i$$

mientras que la varianza de la inversión está dada por

$$\sigma^2 = E[(\sigma - \overline{\sigma})^2] = \sum_{i=1}^n \sum_{j=1}^n x_i \sigma_{ij} x_j$$

Lo que le interesa al inversionista es minimizar la volatilidad de la inversión. Plantee el problema y encuentre la solución.

Solución. El enunciado se traduce en el siguiente problema:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^n \sum_{j=1}^n x_i \sigma_{ij} x_j \\ \text{s.a} \quad & \sum_{i=1}^n r_0^i x_i = y_0, \ \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0 \ \forall i = 1, \dots, n \end{aligned}$$

la última restricción nos sirve para normalizar las cantidades invertidas y si los pesos relativos suman uno, bastará con ponderar la solución óptima por algún escalar y se obtiene la cantidad pedida.

La función lagrangeano para este problema es

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \sum_{i=1}^n \sum_{j=1}^n x_i \sigma_{ij} x_j + \lambda_1 \left(y_0 - \sum_{i=1}^n r_0^i x_i \right) + \lambda_2 \left(1 - \sum_{i=1}^n x_i \right) + \sum_{i=1}^n \mu_i (0 - x_i)$$

Si la solución es interior, es decir $x_i > 0$ $\forall i$ se puede derivar con respecto a x_i y la condición de primer orden nos queda como sigue:

$$\sum_{j=1}^{n} \sigma_{ij} x_j - \lambda_1 r_0^i - \lambda_2 = 0$$

Es posible expresar esta condición considerando todos los activos, esto se obtiene con una expresión vectorial dada por

$$2Q\boldsymbol{x}_0 - \lambda_1 \boldsymbol{e} - \lambda_2 \boldsymbol{r} = \boldsymbol{0}$$

donde Q denota la matriz de covarianzas, $\boldsymbol{e}$ es la suma de los vectores canónicos de $\mathbb{R}^n$ $\boldsymbol{e}_i = (0, \ldots, 1, \ldots, 0)$ y $\boldsymbol{r}_0 = (r_0^1, \ldots, r_0^n)$. Como la solución es interior todos los μ_i son cero. Luego, si $\boldsymbol{e}$ y $\boldsymbol{r}_0$ son linealmente independientes el teorema de Kuhn-Tucker es válido, en otro caso la validez no se pierde porque las restricciones son lineales.

Si Q es invertible, entonces

$$\boldsymbol{x}_0 = \frac{1}{2}\left(Q^{-1}\lambda_1 \boldsymbol{e} + Q^{-1}\lambda_2 \boldsymbol{r}\right)$$

y así se obtendrán las soluciones $(x_0^1, \ldots, x_0^n)$

Por sustitución

$$\boldsymbol{x}_0 \cdot \boldsymbol{e} = 1 \qquad \boldsymbol{x}_0 \cdot \boldsymbol{r}_0 = y_0$$

entonces

$$\begin{aligned} 1 = \boldsymbol{x}_0 \cdot \boldsymbol{e} &= \frac{1}{2}\boldsymbol{e} \cdot (Q^{-1}\boldsymbol{e})\lambda_1 + \frac{1}{2}\boldsymbol{e} \cdot (Q^{-1}\boldsymbol{r}_0)\lambda_2 \\ y_0 = \boldsymbol{x}_0 \cdot \boldsymbol{r} &= \frac{1}{2}\boldsymbol{r}_0 \cdot (Q^{-1}\boldsymbol{e})\lambda_1 + \frac{1}{2}\boldsymbol{r} \cdot (Q^{-1}\boldsymbol{r}_0)\lambda_2 \end{aligned}$$

como esto genera un sistema de ecuaciones se puede resolver para despejar λ_1 y λ_2 que son escalares de la forma $a + bx$ y se obtiene

$$\begin{aligned} \lambda_1 &= a_1 + b_1 y_0 \\ \lambda_2 &= a_2 + b_2 y_0 \end{aligned}$$

donde a y b son constantes que dependen del sistema anterior. Retomando la ecuación

$$\boldsymbol{x}_0 \cdot \boldsymbol{e} = 1 \qquad \boldsymbol{x}_0 \cdot \boldsymbol{r}_0 = y_0$$

si reemplazamos λ_1 y λ_2 se llega a

$$\boldsymbol{x}_0 = y_0 \boldsymbol{d} + \boldsymbol{w}$$

donde $\boldsymbol{d}$ y $\boldsymbol{w}$ son vectores de $\mathbb{R}^n$ que dependen de Q y $\boldsymbol{r}_0$, luego la varianza de la inversión es

$$\sigma^2 = (y_0 \boldsymbol{d} + \boldsymbol{w}) \cdot [Q(y_0 \boldsymbol{d} + \boldsymbol{w})] = (\alpha y_0 + \beta)^2 + \gamma$$

donde α, β y γ dependen de Q y $\boldsymbol{r}_0$.

En base a esto se puede construir una frontera de portafolios eficientes. Cada portafolio eficiente corresponde a un promedio ponderado de dos portafolios que se encuentren en la frontera de eficiencia como se ve en la siguiente figura:

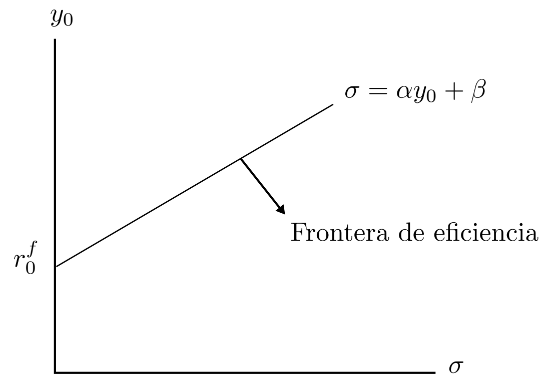


Figura 7.7: Frontera de eficiencia.

$\square$

Podríamos extendernos mucho más sobre optimización pero, de acuerdo a los objetivos del curso, con lo ya expuesto hemos presentado toda la base e incluso muchos más temas que en un curso de cálculo multivariable habitual o estándar. Karush-Kuhn-Tucker nos da la base para la Programación Lineal y muchos métodos eficientes que se tratan en cursos superiores.

7.5. Ejercicios

Teorema de los multiplicadores de Lagrange

Ejercicio 9. Demuestre la siguiente versión del Teorema de los Multiplicadores de Lagrange:

Sea $n \geq 2$ y $D \subseteq \mathbb{R}^n$ abierto. Sean $f, g : D \to \mathbb{R}$ funciones $\mathcal{C}^1$ en D. Supongamos que la restricción de f al conjunto $\Gamma = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) = 0\}$ tiene un punto crítico en $\boldsymbol{a} \in \Gamma$. Entonces $\exists(\lambda, \mu)$ con componentes no nulas tales que

$$\lambda \nabla f(\boldsymbol{a}) + \mu \nabla g(\boldsymbol{a}) = 0 \tag{*}$$

Sugerimos los siguientes pasos:

1. Explique por qué si $\nabla g(\boldsymbol{a}) = 0$ se cumple el teorema.
2. Suponga que $\nabla g(\boldsymbol{a}) \neq 0$. Sin pérdida de generalidad, podemos suponer entonces que $\frac{\partial g(\boldsymbol{a})}{\partial x_n} \neq 0$. Explique la validez de este supuesto.
3. Pruebe que $\exists U \subseteq \mathbb{R}^{n-1}$ que contiene a $\boldsymbol{c} = (a_1, \ldots, a_{n-1})$ y una función $\phi : U \to \mathbb{R}$ tal que $\phi(\boldsymbol{c}) = a_n$ e $(\boldsymbol{y}, \phi(\boldsymbol{y})) \in \Gamma\ \forall \boldsymbol{y} \in U$.
4. Defina $\psi : U \to \mathbb{R}$ por $\psi(\boldsymbol{y}) = f(\boldsymbol{y}, \phi(\boldsymbol{y}))$. Muestre que para $k \in \{1, \ldots, n-1\}$ se cumple que
$$\frac{\partial f(\boldsymbol{a})}{\partial x_k} + \frac{\partial f(\boldsymbol{a})}{\partial x_n} \cdot \frac{\partial \phi(\boldsymbol{c})}{\partial x_k} = 0$$
5. Encuentre una expresión para $\nabla \phi(\boldsymbol{c})$ en función de g o sus derivadas y concluya.

Indicación. Encuentre (λ, μ) tal que la ecuación (*) se cumpla para las primeras $n-1$ coordenadas y verifique y verifique que con esos mismos valores la ecuación también se cumple para la n-ésima coordenada.

Ejercicio 10. Demuestre que si se tienen n números positivos entonces el problema:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \sum_{i=1}^{n} x_i \\ \text{s.a} \quad & \prod_{i=1}^{n} x_i = 1,\ x_i \geq 0\ \forall i \in \{1, \ldots, n\} \end{aligned}$$

Nos permite encontrar la igualdad

$$\frac{1}{n} \sum_{i=1}^{n} x_i \geq \left(\prod_{i=1}^{n} x_i \right)^{1/n}$$

Ejercicio 11. Deduzca que el siguiente problema:

$$\begin{aligned} \max_{x} \quad & \prod_{i=1}^{n} x_i \\ \text{s.a} \quad & \sum_{i=1}^{n} x_i = 1,\ x_i \geq 0\ \forall i \in \{1, \ldots, n\} \end{aligned}$$

Nos permite obtener la misma igualdad del problema anterior.

Notación

$\mathrm{adh}(A)$	Adherencia de A
$B(\boldsymbol{x}_0, r)$	Bola abierta de centro $\boldsymbol{x}_0$ y radio r
$\overline{B}(\boldsymbol{x}_0, r)$	Bola cerrada de centro $\boldsymbol{x}_0$ y radio r
$A^\perp$	Complemento ortogonal de A
$\mathbb{R}_+$	Conjunto de los reales mayores o iguales a cero
$\mathbb{R}_{++}$	Conjunto de los reales estrictamente positivos
$K(\boldsymbol{x})$	Conjunto de direcciones críticas del e.v. S en $\boldsymbol{x}$
$:=$	Definido por
$A \setminus B$	Diferencia entre A y B
Df	Diferencial de f
$\langle \nabla f \rangle$	Espacio vectorial generado por el gradiente de f
$N_S(\boldsymbol{x})$	Espacio vectorial normal al e.v. S en $\boldsymbol{x}$
$T_S(\boldsymbol{x})$	Espacio vectorial tangente al e.v. S en $\boldsymbol{x}$
$\mathcal{C}^n$	Familia de las funciones con n-ésimas derivadas parciales continuas
$\mathrm{fr}(A)$	Frontera de A (puede escribirse ∂A)
∇f	Gradiente de f (vector derivada)
Hf	Hessiano de f
$H_x f$	Hessiano de f sólo con respecto al vector $\boldsymbol{x}$
$\mathrm{int}(A)$	Interior de A
e_j	j-ésimo elemento de la base canónica de $\mathbb{R}^n$
$\triangle f$	Laplaciano de f (sumatoria de todas las segundas derivadas parciales)
$\mapsto$	Mapeo o transformación que describe una función
$\boldsymbol{x} \cdot \boldsymbol{y}$	Producto interno entre $\boldsymbol{x}$ e $\boldsymbol{y}$

Bibliografía

1. Amaya, J. **Optimización Para Estudiantes de Ingeniería**. Departamento de Ingeniería Matemática, Universidad de Chile, 2010.
2. Apostol, T. **Calculus vol II**. Wiley, 1969.
3. Belmar, C., Leiva F. **Curso Completo de Introducción a la Microeconomía**. Departamento de Economía, Universidad de Chile, 2004.
4. Berge, C. **Espaces Topologiques: Fonctions Mulivoques**. Dunod, 1959.
5. Bertsekas, D. **Nonlinear Programming**. Athena Scientific, 1995.
6. Border, K., Aliprantis, C. **Infinite Dimensional Analysis**. Springer, 2006.
7. Border, K. **Fixed Point Theorems with Applications to Economics and Game Theory**. Cambridge University Press, 1989.
8. Boyd, S., Vandenberghe, L. **Convex Optimization**. Cambridge University Press, 2009.
9. Buck, C. **Advanced Calculus**. Mc Graw-Hill, 1965.
10. Cartan, H. **Cálculo Diferencial**. Omega, 1972.
11. Cominetti, R., Jofré A. **Introducción a la Programación Matemática**. IV Simposio Chileno de Matemática, 1993.
12. Fleming, W. **Funciones de Varias Variables**. CECSA, 1969.
13. Jehle, G. **Advanced Microeconomic Theory**. Addison Wesley, 2000.
14. Karush, W. **Minima of Functions of Several Variables with Inequalities as Side Constraints**. M. Sc. Dissertation, Dept. of Mathematics, University of Chicago, 1939.
15. Kolmogorov, A., Fomin, S. **Introductory Real Analysis**. Dover Publications, 1975.
16. Kopka, H., Daly, P. **Guide to LaTeX**. Addison-Wesley, 2003.
17. Krell, R., Rojas, E., Torres-Martínez, J. **Apuntes de Microeconomía I: Teoría del Consumidor**. Departamento de Economía, Universidad de Chile, 2009.
18. Kuhn, H., Tucker, W. **Nonlinear Programming**. Proceedings of 2nd Berkeley Symposium, University of California Press, 1951.
19. Luenberger, D., Ye, Y. **Linear and Nonlinear Programming**. Springer, 2008.

20. Mathworks Team. **Learning Matlab 6 Student Version**. Mathworks, 2001.
21. Mas Colell, A., Whinston, D., Green, J. **Microeconomic Theory**. Oxford University Press, 1995.
22. Nocedal, J., Wright S. **Numerical Optimization**. Springer, 1999.
23. Ok, E. **Real Analysis with Economic Applications**. Princeton University Press, 2007.
24. Ortiz, C., Varas, S., Vera, J. **Optimización y Modelos Para la Gestión**. Dolmen Ediciones, 2000.
25. Piskunov, N. **Cálculo Diferencial e Integral, (vol. I y II)**. Editorial MIR, 1970.
26. Rivera, J. **Microeconomía II: Apunte de Curso**. Departamento de Economía, Universidad de Chile, 2008.
27. Rosenlicht, M. **Introduction to Analysis**. Dover Publications, 1985.
28. Royden, H. **Real Analysis**. Macmillan Company, 1968.
29. Ruszczyński, A. **Nonlinear Optimization**. Princeton University Press, 2006.
30. Simon, C., Blume, L. **Mathematics for Economists**. W.W. Norton & Company, 1994.
31. Spivak, M. **Calculus**. Publish or Perish, 2008.
32. Stewart, J. **Cálculo Multivariable**. Thomson Learning, 2006.
33. Takayama, A. **Mathematical Economics**. Cambridge University Press, 1985.
34. Torres-Martínez, J.P. **Apuntes de Matemática Para Microeconomía Avanzada**. Departamento de Economía, Universidad de Chile, 2010.
35. Varian, H. **Microeconomic Analysis**. W.W Norton & Company, 1992.

Índice alfabético