Apuntes de la Escuela de Ingeniería

ING2001 - Cálculo Avanzado

Gianfranco Liberona

David Salas

Andrés Zúñiga

Universidad de O'Higgins

Primavera-2024

Aclaraciones sobre el apunte y agradecimientos

Este documento fue confeccionado como material de estudio para el curso de Cálculo Avanzado ING2001 de la Universidad de O'Higgins, sin fines de lucro. Está basado sobre los apuntes del curso dictado por los profesores Gianfranco Liberona Henríquez, David Salas Videla y Andrés Zúñiga Munizaga, entre los años 2020 y 2024. Se agradece la colaboración de Camila Romero, como editora de la presente versión, y a la Dirección de Pregrado, que apoyó la confección de este documento a través de los Fondos de Innovación en Docencia.

Índice general

Capítulo 1

Geometría y Topología en $\mathbb{R}^n$

1.1 El Espacio $\mathbb{R}^n$

En el curso de álgebra lineal, estudiamos el espacio vectorial $\mathbb{R}^n$ (con $n \in \mathbb{N}$), que consiste en el conjunto de vectores columna de n entradas a valores reales.

Nota: Conjuntos numéricos

En este curso ocuparemos la siguiente notación:

- $\mathbb{N} = \{1, 2, 3, \cdots\} \to$ Números naturales.
- Para todo $n \in \mathbb{N}$, tomamos $[n] = \{1, 2, \ldots, n\}$. Es decir, $[n]$ es el conjunto finito de los primeros n números naturales.
- $\mathbb{Z} = \{\cdots, -1, 0, 1, \cdots\} \to$ Números enteros.
- $\mathbb{Q} \to$ Números racionales.
- $\mathbb{R} \to$ Números Reales.

Observar que para nosotros, $\mathbb{N}$ **no contiene al cero**. Esto es una convención. Otros libros consideran al cero como número natural.

Un ejemplo tradicional de espacio vectorial es $\mathbb{R}^3$ (consiste en tomar $n = 3$) y se interpreta cada coordenada como la posición en el espacio con respecto a un punto de referencia, el origen $O = (0, 0, 0)$. Es decir, utilizamos un vector $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ para describir la posición de un punto en el espacio, siguiendo la asignación:

$$\begin{aligned} \text{Largo} &\to x_1, \\ \text{Ancho} &\to x_2, \\ \text{Alto} &\to x_3. \end{aligned}$$

Más aún, podemos representar gráficamente el punto descrito, usando el espacio cartesiano descrito por el sistema de referencia de tres rectas ortogonales: los ejes X_1, X_2 y X_3, también denotados como ejes X, Y y Z (ver Figura 1.1).

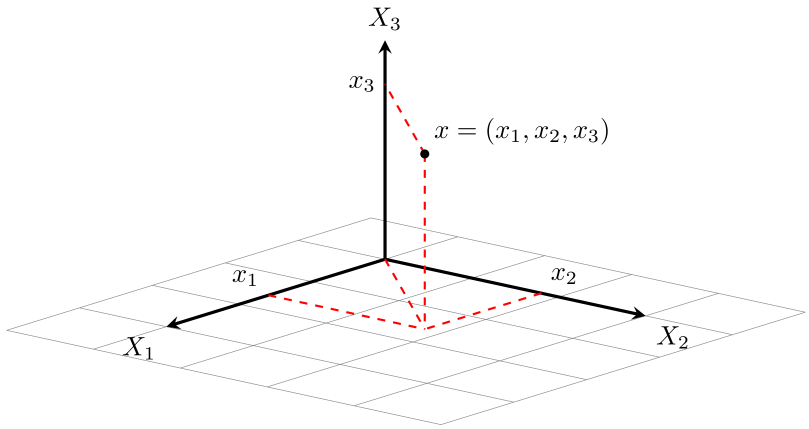


Figura 1.1: Representación gráfica de $\mathbb{R}^3$. Cada vector $x = (x_1, x_2, x_3)$ está descrito por sus proyecciones a los ejes de coordenadas X_1, X_2 y X_3.

En el caso general, cuando n puede ser cualquier número natural, el espacio $\mathbb{R}^n$ nos sirve para describir objetos con n características numéricas, siguiendo el esquema:

$$\begin{aligned}
\text{Característica 1} &\quad\rightarrow\quad x_1,\\
\text{Característica 2} &\quad\rightarrow\quad x_2,\\
\text{Característica 3} &\quad\rightarrow\quad x_3,\\
&\quad\vdots\\
\text{Característica n} &\quad\rightarrow\quad x_n.
\end{aligned}$$

Así, objetos complejos, como las baterías eléctricas o los mercados de acciones, pueden ser descritos usando vectores en $\mathbb{R}^n$: cada vector $x = (x_1, \ldots, x_n)$ describe una posible configuración de los objetos modelados.

1.1.1 $\mathbb{R}^n$ como espacio vectorial

Recordemos la definición de $\mathbb{R}^n$ como el conjunto de todos los vectores de n coordenadas.

Definición 1.1.1. *Sea $n \in \mathbb{N}$ fijo. Diremos que x es un vector de n coordenadas si es una matriz de n filas y una columna con valores en $\mathbb{R}$, es decir, si es de la forma*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad \textit{con } x_1, x_2, \ldots, x_n \in \mathbb{R}. \tag{1.1}$$

Denotaremos alternativamente el vector x de entradas $x_1, \ldots, x_n$, como $x = (x_1, \ldots, x_n)$, $x = (x_i)_{i=1}^n$ o también $(x_i \ : \ i \in [n])$. Definimos $\mathbb{R}^n$ como el conjunto de todos los vectores de n coordenadas, es decir,

$$\mathbb{R}^n = \left\{ x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \, x_1, \ldots, x_n \in \mathbb{R} \right\}. \tag{1.2}$$

Atención

En este apunte, ocuparemos indistintamente la notación de n-tupla y la notación vectorial para describir vectores en $\mathbb{R}^n$, es decir,

$$x = (x_1, \ldots, x_n) \iff x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Sin embargo, al momento de operar vectores, ocuparemos la convención de que los vectores en $\mathbb{R}^n$ **siempre son vectores columna**. La notación de n-tupla la ocuparemos sobre todo para ahorrar espacio en texto.

El espacio $\mathbb{R}^n$ posee una estructura vectorial natural heredada de las operaciones de suma y multiplicación de los números reales.

Definición 1.1.2. *Sean $n \in \mathbb{N}, x = (x_i)_{i=1}^n \in \mathbb{R}^n, y = (y_i)_{i=1}^n \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$. Definimos:*

1. *La **suma** de x e y como el vector*

$$x + y = (x_i + y_i)_{i=1}^n := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \tag{1.3}$$

2. *La **ponderación** de x por λ como el vector*

$$\lambda x = (\lambda x_i)_{i=1}^n := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \tag{1.4}$$

Con estas operaciones, $(\mathbb{R}^n, +, \cdot)$ es un espacio vectorial de dimensión n.

En el curso de álgebra lineal se estudiaron varias propiedades algebraicas y geométricas del espacio vectorial $(\mathbb{R}^n, +, \cdot)$.

1.1.2 Norma y distancia en $\mathbb{R}^n$

Dados dos vectores $x, y \in \mathbb{R}^n$, es natural preguntarse qué tan diferentes son el uno del otro. Esta pregunta es particularmente pertinente en problemas de modelamiento, donde x e y pueden ser

- resultados diferentes de un mismo experimento;
- resultados numéricos sujetos a errores computacionales;
- descripciones de dos objetos de un mismo modelo, sujetos a variaciones en el proceso de producción;
- etcétera.

Saber qué tan diferentes (o parecidos) son dos vectores, nos permite evaluar cosas como la replicabilidad de un experimento, la certitud de un proceso computacional o los estándares de producción de un producto.

Una alternativa para medir la diferencia de dos vectores es usar alguna noción de *distancia*. En el caso de $\mathbb{R}$, somos capaces de medir la distancia entre dos números $x, y \in \mathbb{R}$ usando el valor absoluto $|\cdot|$. Es decir, definimos la distancia de x a y en la recta real como

$$d(x,y) := |x-y|. \tag{1.5}$$

En $\mathbb{R}^n$, definiremos la distancia entre dos vectores de la misma manera, usando una generalización del valor absoluto llamada *norma euclidiana*, o simplemente norma.

Definición 1.1.3 (Norma Euclidiana)**.** *Sea $n \in \mathbb{N}$. Para un vector $x \in \mathbb{R}^n$ definimos la **norma** de x como*

$$\|x\| := \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}. \tag{1.6}$$

Cuando $n = 1$, la norma $\|x\|$ coincide con el valor absoluto, por lo que escribimos simplemente $|x|$.

Atención

En rigor, para cada valor $n \in \mathbb{N}$, la norma define una función diferente $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$, que deberíamos denotar de manera diferente cada vez.
Sin embargo, para evitar recargar la notación, se ocupa la misma notación $\|\cdot\|$ para todos los valores de la dimensión $n \in \mathbb{N}$, entendiéndose por contexto a cual nos referimos.

La norma euclidiana nos da una noción de **tamaño** de un vector x en $\mathbb{R}^n$, midiendo el largo del segmento que va del origen 0 al punto $x = (x_1, x_2, \ldots, x_n)$. Recibe el nombre de norma euclidiana pues el "largo" del segmento se calcula usando el teorema de Pitágoras de la geometría Euclidiana.

Proposición 1.1.4. *Sea $n \in \mathbb{N}$. La norma euclidana $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ cumple las siguientes propiedades:*

1. ***(Positividad)*** *Para todo $x \in \mathbb{R}^n, \|x\| \geq 0$.*
2. ***(Separación)*** *Para todo $x \in \mathbb{R}^n, \|x\| = 0 \iff x = 0$.*
3. ***(Homogeneidad)*** *Para todo $x \in \mathbb{R}^n$ y todo $\alpha \in \mathbb{R}, \|\alpha x\| = |\alpha|\|x\|$.*
4. ***(Desigualdad Triangular)*** *Para todo $x, y \in \mathbb{R}^n, \|x + y\| \leq \|x\| + \|y\|$.*

Demostración. Probaremos cada afirmación por separado:

1. Positividad: La suma de números al cuadrado es siempre positiva, recordando que la raíz cuadrada de un número positivo es positiva, se puede concluir que para cualquier $x \in \mathbb{R}^n$:

$$\|x\| = (\underbrace{x_1^2 + \cdots + x_n^2}_{\geq 0})^{1/2} \geq 0.$$

2. Separación: Para esta demostración se usan dos propiedades de Cálculo en $\mathbb{R}$:

- La unicidad de solución de la ecuación $\alpha^2 = 0$ (*El único número que al cuadrado es cero, es el cero*):
$$\forall \alpha \in \mathbb{R}, \quad \alpha^2 = 0 \iff \alpha = 0 \tag{1.7}$$
- Una suma de números positivos es cero si y solo si todos los sumandos son cero:
$$\sum_{i=1}^{n} x_i = 0 \quad \iff \quad \forall i \in [n],\, x_i = 0 \tag{1.8}$$

Utilizando lo anterior podemos escribir:

$$\begin{aligned}
\|x\| = 0 &\iff \|x\|^2 = 0 && \text{(1.7) con } \alpha = \|x\| \\
&\iff \sum_{i=1}^{n} x_i^2 = 0 && \text{definición de } \|x\| \\
&\iff \forall i \in [n],\ x_i^2 = 0 && \text{(1.8)} \\
&\iff \forall i \in [n],\ x_i = 0 && \text{(1.7) con } \alpha = x_i \\
&\iff x = (x_1, \ldots, x_n) = 0.
\end{aligned}$$

3. Homogeneidad: Sea $\alpha \in \mathbb{R}$. Ocupando la definición de la ponderación dada por (1.4), podemos escribir

$$\|\alpha x\| = \sqrt{\sum_{i=1}^{n} (\alpha x_i)^2} = \sqrt{\sum_{i=1}^{n} \alpha^2 x_i^2} = \sqrt{\alpha^2 \sum_{i=1}^{n} x_i^2} = \sqrt{\alpha^2}\sqrt{\sum_{i=1}^{n} x_i^2} = |\alpha|\|x\|.$$

4. Desigualdad Triangular: Demostraremos el resultado por inducción en la dimensión n, usando dos casos bases.

1. **Caso base, $n = 1$:** Tomemos $x, y \in \mathbb{R}$, queremos mostrar que $|x + y| \leq |x| + |y|$.

   Sumando las desigualdades $x \leq |x|$ e $y \leq |y|$ por un lado, y las desigualdades $-x \leq |x|$ e $-y \leq |y|$, obtenemos que

   $$x + y \leq |x| + |y| \quad \text{y} \quad -x - y = -(x + y) \leq |x| + |y|.$$

   Luego,

   $$|x + y| = \text{máx}\{x + y, -(x + y)\} \leq |x| + |y|.$$

2. **Caso base, $n = 2$:** Para esta parte usaremos la monotonía de la función raíz cuadrada, o en otras palabras, que tomar raiz mantiene las desigualdades:

   $$\forall a, b \in \mathbb{R}_+, \qquad a < b \quad \implies \quad \sqrt{a} < \sqrt{b}. \tag{1.9}$$

   También es importante recordar que la raíz cuadrada de un número cuadrado da **el valor absoluto** de dicho número ($\sqrt{a^2} = |a|$). En el caso de la norma, **dada su positividad**, ocurre que $\sqrt{\|x\|^2} = \|x\|$. Por todo lo anterior, en vez de trabajar la norma, utilizaremos la norma al cuadrado.

   Tomemos $x, y \in \mathbb{R}^2$. Luego, podemos escribir

   $$\begin{aligned}
   \|x + y\|^2 &= (x_1 + y_1)^2 + (x_2 + y_2)^2 \\
   &= x_1^2 + 2x_1y_1 + y_1^2 + x_2^2 + 2x_2y_2 + y_2^2 \\
   &= (x_1^2 + y_1^2) + 2\underline{(x_1y_1 + x_2y_2)} + (x_2^2 + y_2^2).
   \end{aligned}$$

   Ahora vamos a acotar el término subrayado, al cuadrado por el mismo argumento anterior de (1.9). Además utilizaremos la desigualdad de cuadrado de binomio, donde para todo $a, b \in \mathbb{R}$ se tiene que $2ab \leq a^2 + b^2$.

   Entonces podemos escribir lo siguiente:

   $$\begin{aligned}
   (x_1y_1 + x_2y_2)^2 &= x_1^2y_1^2 + 2x_1y_1x_2y_2 + x_2^2y_2^2 \\
   &= x_1^2y_1^2 + 2(x_1y_2)(x_2y_1) + x_2^2y_2^2 \\
   &\leq x_1^2y_1^2 + x_1^2y_2^2 + x_2^2y_1^2 + x_2^2y_2^2 \\
   &= (x_1^2 + x_2^2)(y_1^2 + y_2^2) = \|x\|^2\|y\|^2. \\
   \implies |x_1y_1 + x_2y_2| &\leq \|x\|\|y\| \\
   \implies (x_1y_1 + x_2y_2) &\leq \|x\|\|y\|
   \end{aligned}$$

Retomando el desarrollo anterior, tenemos que

$$\begin{aligned}
\|x+y\|^2 &= (x_1^2+y_1^2)+2(x_1y_1+x_2y_2)+(x_2^2+y_2^2)\\
&\leq \|x\|^2+2\|x\|\|y\|+\|y\|^2\\
&= (\|x\|+\|y\|)^2\\
\implies \|x+y\| &\leq \|x\|+\|y\|
\end{aligned}$$

Notemos que al tomar raíz de $(\|x\|+\|y\|)^2$, no es necesario el valor absoluto porque $(\|x\|+\|y\|)$ es positivo.

3. **Paso inductivo** $n \implies n+1$:

   Tomemos $n \geq 2$ y supongamos que la desigualdad triangular se cumple en $\mathbb{R}^n$. Tomemos ahora $x, y \in \mathbb{R}^{n+1}$ y definamos $\hat{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ y $\hat{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Podemos escribir, usando la hipótesis inductiva y el primer caso base, que

$$\begin{aligned}
\|x+y\|^2 = \sum_{i=1}^{n+1}(x_i+y_i)^2 &= \left(\sum_{i=1}^{n}(x_i+y_i)^2\right)+(x_{n+1}+y_{n+1})^2\\
&= \|\hat{x}+\hat{y}\|^2+|x_{n+1}+y_{n+1}|^2\\
&\leq \|\hat{x}\|^2+\|\hat{y}\|^2+|x_{n+1}|^2+|y_{n+1}|^2\\
&\leq (\|\hat{x}\|+\|\hat{y}\|)^2+(|x_{n+1}|+|y_{n+1}|)^2.
\end{aligned}$$

   Definamos ahora $u = (\|\hat{x}\|, |x_{n+1}|)$ y $v = (\|\hat{y}\|, |y_{n+1}|)$, ambos vectores en $\mathbb{R}^2$. Notemos que

$$\begin{aligned}
\|u\| &= \sqrt{\|\hat{x}\|^2+|x_{n+1}|^2} = \sqrt{\left(\sum_{i=1}^{n}x_i^2\right)+x_{n+1}^2} = \|x\|^2,\\
\|v\| &= \sqrt{\|\hat{y}\|^2+|y_{n+1}|^2} = \sqrt{\left(\sum_{i=1}^{n}y_i^2\right)+y_{n+1}^2} = \|y\|^2.
\end{aligned}$$

   Retomando la desigualdad anterior y usando el segundo caso base, tenemos que

$$\begin{aligned}
\|x+y\|^2 &\leq (\|\hat{x}\|+\|\hat{y}\|)^2+(|x_{n+1}|+|y_{n+1}|)^2\\
&= \|u+v\|^2\\
&\leq (\|u\|+\|v\|)^2\\
&= (\|x\|+\|y\|)^2.
\end{aligned}$$

   Tomando raíz cuadrada a ambos lados de la ecuación, se deduce la desigualdad triangular para los vectores $x, y \in \mathbb{R}^{n+1}$.

Por inducción, tenemos que la desigualdad triangular se cumple para cualquier dimensión n, y para cualquier par de vectores $x, y \in \mathbb{R}^n$. Esto concluye la demostración. $\square$

La desigualdad triangular es una de las propiedades más importantes de la norma euclidiana. Si bien la demostración es puramente analítica, también es posible interpretar el resultado de manera geométrica, usando la regla del paralelogramo de la suma de dos vectores en $\mathbb{R}^3$. La idea es la siguiente:

- Dados dos vectores $x, y \in \mathbb{R}^n$ no paralelos, podemos mirarlos como dos vectores en $H = \text{Gen}\{x, y\}$.

- Al ser no-paralelos, los vectores x e y deben ser linealmente independientes y por lo tanto $\dim(H) = 2$. Es decir, H tiene la misma geometría que $\mathbb{R}^2$.
- Podemos mirar en H el triángulo que forman x, y y $x + y$, usando la siguiente regla de lados de un triángulo: El tamaño de un lado de un triángulo debe ser menor que la suma de los otros dos lados.
- Con la observación anterior, tenemos que el tamaño del segmento que une 0 con $x + y$ debe ser menor que la suma de los segmentos que unen 0 con x, y x con $x + y$.

La Figura 1.2 ilustra este enfoque en $\mathbb{R}^3$. La idea es que al tener solo dos vectores $x, y \in \mathbb{R}^n$ involucrados, podemos reducir nuestro análisis al caso de $\mathbb{R}^2$, usando el espacio $H = \text{Gen}\{x, y\}$.

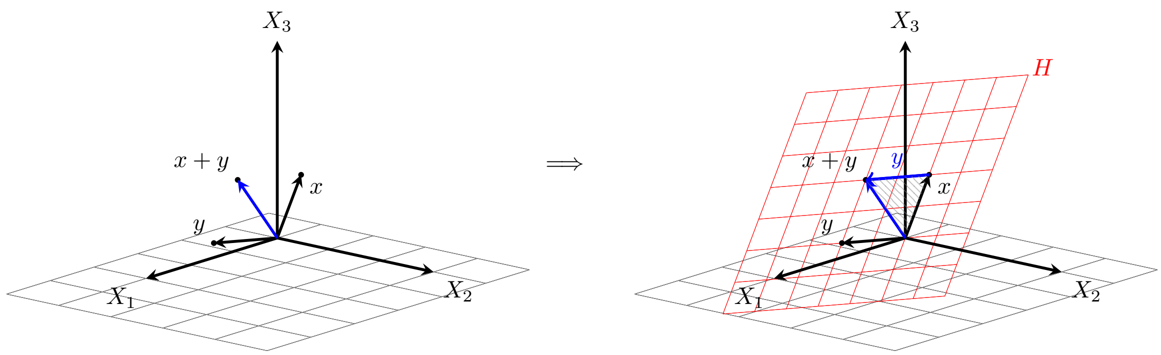


Figura 1.2: Ilustración de la desigualdad triangular en $\mathbb{R}^3$. El espacio $H = \text{Gen}\{x, y\}$ se ilustra con la grilla de color rojo. Los segmentos $0 \to x$, $x \to (x + y)$ y $0 \to (x + y)$ forman un triángulo en H. El segmento $x \to (x + y)$ coincide con el vector y, partiendo de x.

Nota: Desigualdad de Cauchy-Schwartz en $\mathbb{R}^2$

En la demostración de la desigualdad triangular, demostramos como paso intermedio la siguiente desigualdad auxiliar en $\mathbb{R}^2$:

$$\forall x, y \in \mathbb{R}^2,\ (x_1y_1 + x_2y_2)^2 \leq (x_1^2 + x_2^2)(y_1^2 + y_2^2). \tag{1.10}$$

En realidad, esta desigualdad es posible de generalizarla en dimensión arbitraria, y se conoce como la desigualdad de Cauchy-Schwartz. El caso general lo estudiaremos en la Sección 1.1.3, particularmente en el Coroloario 1.1.12.

Habiendo definido la norma como el tamaño de un vector, podemos definir la *distancia* entre dos vectores como el tamaño de su diferencia.

Definición 1.1.5. *Sea $n \in \mathbb{N}$, y sean dos vectores $x, y \in \mathbb{R}^n$. Definimos la* ***distancia euclidiana*** *entre x e y como*

$$d(x, y) := \|x - y\|. \tag{1.11}$$

La función distancia $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ hereda sus propiedades de la norma.

Proposición 1.1.6. *Sea $n \in \mathbb{N}$. La función distancia $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ cumple las siguientes propiedades:*

1. ***(Positividad)*** *Para todo $x, y \in \mathbb{R}^n, d(x, y) \geq 0$.*
2. ***(Simetría)*** *Para todo $x, y \in \mathbb{R}^n, d(x, y) = d(y, x)$.*

3. ***(Separación)*** *Para todo* $x, y \in \mathbb{R}^n, d(x, y) = 0 \iff x = y$.

4. ***(Invarianza bajo traslación)*** *Para todo* $x, y, z \in \mathbb{R}^n, d(x, y) = d(x + z, y + z)$.

5. ***(Desigualdad Triangular)*** *Para todo* $x, y, z \in \mathbb{R}^n, d(x, y) \leq d(x, z) + d(z, y)$.

Demostración. Las propiedades de positividad, simetría, separación e invarianza bajo traslación son consecuencias directas de las propiedades de la norma. En efecto,

- $d(x, y) = \|x - y\| \geq 0$ por positividad de la norma, Proposición 1.1.4.
- $d(x, y) = \|x - y\| = \|y - x\| = d(y, x)$, por la Definición 1.1.3.
- $d(x, y) = 0 \iff \|x - y\| = 0 \iff x - y = 0 \iff x = y$, por separación de la norma, Proposición 1.1.4.
- $d(x, y) = \|x - y\| = \|x + z - y - z\| = d(x + z, y + z)$, por la Definición 1.1.3.

Finalmente, para la desigualdad triangular consideramos el siguiente desarrollo:

$$\begin{aligned}
d(x, y) &= d(x - z, y - z) && \text{(Invarianza bajo traslación)} \\
&= \|(x - z) - (y - z)\| && \text{(definición distancia)} \\
&= \|(x - z) + (z - y)\| \\
&\leq \|x - z\| + \|z - y\| && \text{(Desigualdad triangular norma)} \\
&= d(x, z) + d(z, y)
\end{aligned}$$

□

1.1.3 Producto escalar

La geometría euclidiana está fuertemente ligada a lo que se conoce como el *producto escalar* entre vectores, y su relación con los ángulos que se forman entre ellos. Comencemos recordando la definición.

Definición 1.1.7. *Sean* $x, y \in \mathbb{R}^n$. *Definimos el producto escalar entre* x *e* y *como*

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

Dependiendo del autor, el producto escalar también es llamado producto interno o producto punto, y alternativamente se escribe como $\langle x, y \rangle = x \cdot y$.

Proposición 1.1.8. *Se tienen las siguientes propiedades:*

1. ***(Relación con norma Euclidiana)*** *Para todo* $x \in \mathbb{R}^n$, $\langle x, x \rangle = \|x\|^2$.

2. ***(Absorción del cero)*** *Para todo* $x \in \mathbb{R}^n$, $\langle 0, x \rangle = 0$.

3. ***(Separación)*** *Se tiene que* $\langle x, x \rangle = 0 \iff x = 0$.

4. ***(Simetría)*** *Para todo* $x, y \in \mathbb{R}^n$, $\langle x, y \rangle = \langle y, x \rangle$.

5. ***(Bilinealidad)*** *Para todo* $x, y, z \in \mathbb{R}^n$ *y todo* $\lambda \in \mathbb{R}$, *se tiene que* $\langle x, y + \lambda z \rangle = \langle x, y \rangle + \lambda \langle x, z \rangle$

Demostración. Demostraremos cada afirmación por separado.

1. Relación con Norma Euclideana: Sea $x \in \mathbb{R}^n$. Tenemos que

$$\langle x, x \rangle = \sum_{i=1}^{n} x_i x_i = \sum_{i=1}^{n} x_i^2 = \|x\|^2.$$

2. Absorción del cero: Sea $x \in \mathbb{R}^n$. Utilizando la absorción del cero en la multiplicación en $\mathbb{R}$, tenemos que

$$\langle 0, x \rangle = \sum_{i=1}^{n} 0 \cdot x_i = \sum_{i=1}^{n} 0 = 0.$$

3. Separación: Sea $x \in \mathbb{R}^n$. Recordando que $\|x\| = 0 \iff x = 0$ (ver Proposición 1.1.4), podemos escribir

$$\langle x, x \rangle = 0 \iff \|x\| = 0 \iff x = 0.$$

4. Simetría: Sea $x \in \mathbb{R}^n$. Utilizando la conmutatividad en la multiplicación en $\mathbb{R}$, tenemos que

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

5. Biliniealidad: Sean $x, y, z \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$. Se tiene que

$$\begin{aligned}
\langle x, y + \lambda z \rangle &= \sum_{i=1}^{n} x_i (y_i + \lambda z_i) \\
&= \sum_{i=1}^{n} (x_i y_i + \lambda x_i z_i) \\
&= \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} \lambda x_i z_i = \sum_{i=1}^{n} x_i y_i + \lambda \sum_{i=1}^{n} x_i z_i = \langle x, y \rangle + \lambda \langle x, z \rangle.
\end{aligned}$$

La demostración está completa. $\square$

Nota: Bilinealidad del producto escalar

Consideremos la **función** producto escalar

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (x, y) &\mapsto \langle x, y \rangle. \end{aligned}$$

La bilinealidad (propiedad 5 de la Proposición 1.1.8) nos dice que el producto escalar **no es una función lineal**. Por ejemplo, al evaluar la ponderación por escalar, para ser lineal necesitamos que:

$$\langle \lambda(x, y) \rangle = \lambda \langle x, y \rangle$$

Sin embargo,

$$\begin{aligned} \langle \lambda(x, y) \rangle &= \langle \lambda x, \lambda y \rangle \\ &= \lambda \langle x, \lambda y \rangle \\ &= \lambda^2 \langle x, y \rangle \\ &\neq \lambda \langle x, y \rangle \end{aligned}$$

Ahora bien, cuando fijamos una de las dos coordenadas del producto escalar, la función resultante sí es lineal. Dicho de otro modo, para todo $x \in \mathbb{R}^n$, tenemos que las funciones "multiplicar por x"

$$\begin{aligned} \langle x, \cdot \rangle : \mathbb{R}^n &\to \mathbb{R} \\ y &\mapsto \langle x, y \rangle, \end{aligned} \qquad \text{y} \qquad \begin{aligned} \langle \cdot, x \rangle : \mathbb{R}^n &\to \mathbb{R} \\ y &\mapsto \langle y, x \rangle, \end{aligned}$$

son ambas lineales. La linealidad de $\langle x, \cdot \rangle$ es consecuencia de la propiedad 5 de la Proposición 1.1.8. La linealidad de $\langle \cdot, x \rangle$ se deduce de la simetría (propiedad 4 de la Proposición 1.1.8): de hecho la simetría nos permite concluir que las funciones $\langle x, \cdot \rangle$ y $\langle \cdot, x \rangle$ son iguales.

De la relación entre producto escalar y norma euclidiana, se deduce lo que se conoce como ley del paralelogramo que, para dos vectores u y v en $\mathbb{R}^n$, nos permite calcular $\|u + v\|$ y $\|u - v\|$. La razón del nombre es que, cuando u y v son linealmente independientes, los vectores $u + v$ y $u - v$ son las diagonales del paralelogramo formado por los vectores u y v, como se puede ver en la Figura 1.3

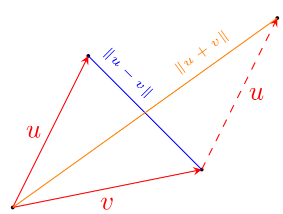


Figura 1.3: Ilustración de la Regla del paralelogramo en $\mathbb{R}^2$

Proposición 1.1.9 (Ley del paralelogramo)**.** *Sea $n \in \mathbb{N}$ y sean $u, v \in \mathbb{R}^n$. Se tiene que*

$$\|u + v\|^2 = \|u\|^2 + 2\langle u, v \rangle + \|v\|^2. \tag{1.12}$$

$$\|u - v\|^2 = \|u\|^2 - 2\langle u, v \rangle + \|v\|^2. \tag{1.13}$$

Demostración. Sean $u, v \in \mathbb{R}^n$. Usando las propiedades de la Proposición 1.1.3 podemos escribir

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle && \text{(Relación con la norma)}\\
&= \langle u+v, u\rangle + \langle u+v, v\rangle && \text{(Bilinealidad)}\\
&= \langle u, u\rangle + \langle v, u\rangle + \langle u, v\rangle + \langle v, v\rangle && \text{(Bilinealidad)}\\
&= \langle u, u\rangle + 2\langle u, v\rangle + \langle v, v\rangle && \text{(Simetría)}\\
&= \|u\|^2 + 2\langle u, v\rangle + \|v\|^2. && \text{(Relación con la norma)}
\end{aligned}$$

Esto demuestra la relación (1.12). La demostración de (1.13) es equivalente, reemplazando v por $(-v)$ en el desarrollo anterior, y notando que $2\langle u, -v\rangle = -2\langle u, v\rangle$. $\square$

El producto interno entre dos vectores está íntimamente ligado a la proyección o *sombra* que produce uno sobre el otro. Para esto, recordaremos la noción de ángulo entre dos vectores.

Definición 1.1.10 (Ángulo entre vectores)**.** *Sea $n \in \mathbb{N}$ con $n \geq 2$ y sean dos vectores u y v en $\mathbb{R}^n$, distintos de 0. Definimos el ángulo entre el vector u y el vector v, cómo el ángulo más pequeño formado por los segmentos $0 \to u$ y $0 \to v$, en el plano $H = \mathrm{Gen}\{u, v\}$.*

Nota: Ángulos en $\mathbb{R}^n$

Si bien $u, v \in \mathbb{R}^n$, su generado tiene dimesión a lo más 2 ($\dim H \leq 2$), por lo que se puede reducir el análisis al plano de $\mathbb{R}^2$. Con esa consideración, el ángulo entre dos vectores u y v es el ángulo en 0 del triángulo formado por los puntos u, v y 0.
En caso que $u = \alpha v$, entonces el ángulo es 0 si $\alpha > 0$ (mismo sentido), o π si $\alpha < 0$ (sentido opuesto).
Notar que por definición, el ángulo entre dos vectores toma valores entre 0 y π.

Denotando θ al ángulo entre u y v, podemos definir la sombra de u sobre la recta generada por v como la cantidad $r = \|u\|\cos(\theta)$. Esta cantidad se construye considerando el único triángulo rectángulo que tiene como hipotenusa el vector u y uno de sus catetos sobre la recta generada por v. La Figura 1.4 ilustra esta construcción.

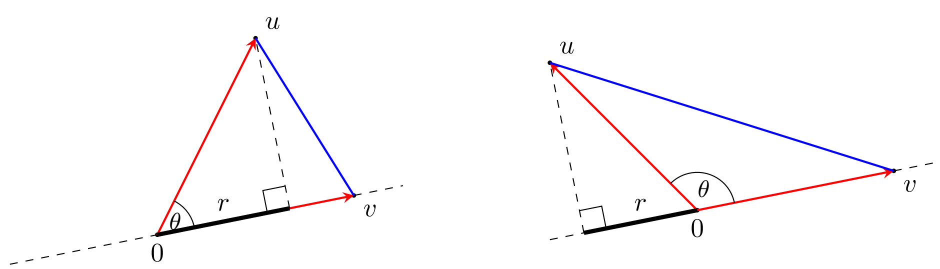


Figura 1.4: Ilustración de sombra del vector u sobre el vector v. A la izquierda, el caso en que el ángulo entre u y v es agudo (entre 0 y $\pi/2$). A la derecha, el caso en que el ángulo entre u y v es obtuso (entre $\pi/2$ y π). En ambos casos, $r = \|u\|\cos(\theta)$.

Teorema 1.1.11. *Sea $n \in \mathbb{N}$ con $n \geq 2$, y sean $u, v \in \mathbb{R}^n$ distintos de cero. Sea θ el ángulo entre u y v. Se tiene que*

$$\langle u, v\rangle = \|u\|\|v\|\cos(\theta) \tag{1.14}$$

Demostración. Fijemos $n \geq 2$. Sin perder generalidad, consideremos que $\|u\| \leq \|v\|$. En caso contrario, bastaría intercambiar los roles de u y v en este desarrollo. Estudiaremos 5 casos, dependiendo del valor del ángulo θ.

Caso 1, $\theta = 0$: entonces $u = \alpha v$ con $\alpha > 0$ y también $\cos(0) = 1$ por lo tanto,

$$\langle u, v\rangle = \langle \alpha v, v\rangle = \alpha\|v\|^2 = \|\alpha v\|\|v\| = \|u\|\|v\| = \|u\|\|v\|\cos(\theta).$$

Caso 2, $\theta = \pi$: entonces $u = \alpha v$ con $\alpha < 0$ y también $\cos(\theta) = -1$ por lo tanto,

$$\langle u, v\rangle = \langle \alpha v, v\rangle = \alpha\|v\|^2 = -|\alpha|\|v\|^2 = -\|\alpha v\|\|v\| = -\|u\|\|v\| = \|u\|\|v\|\cos(\theta).$$

Caso 3, $\theta = \pi/2$: entonces podemos aplicar el teorema de Pitágoras al triángulo rectángulo formado por u y v, como se ilustra en la Figura 1.5.

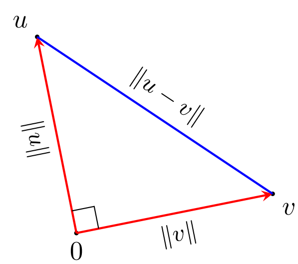


Figura 1.5: Triangulo rectángulo formado por u y v.

Podemos entonces escribir

$$\|u\|^2 + \|v\|^2 = \|u - v\|^2.$$

Por otro lado, gracias a la Ley del paralelogramo (Proposición 1.1.9), sabemos que

$$\|u - v\|^2 = \|u\|^2 - 2\langle u, v\rangle + \|v\|^2.$$

Juntando ambas ecuaciones, concluimos que $\langle u, v\rangle = 0$. Como $\cos(\pi/2) = 0$, se cumple el resultado.

Caso 4, $\theta \in (0, \frac{\pi}{2})$: Ahora, el ángulo entre u y v es agudo. Entonces, podemos dibujar la Figura 1.6, inducida por el triángulo formado por u y v. De aquí podemos escribir las siguientes ecuaciones:

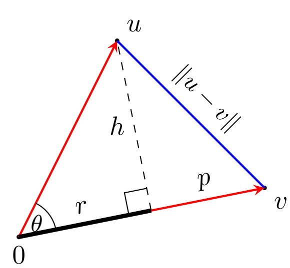


Figura 1.6: Triangulo formado por u y v.

$$\begin{aligned} r^2 + h^2 &= \|u\|^2, \\ h^2 + p^2 &= \|u - v\|^2, \\ r + p &= \|v\|. \end{aligned}$$

Despejando $h^2 = \|u - v\|^2 - p^2$ y usando la regla del paralelogramo (1.13), podemos escribir

$$\begin{aligned} r^2 &= \|u\|^2 - h^2 \\ &= \|u\|^2 - \|u - v\|^2 + p^2 \\ &= \|u\|^2 - (\|u\|^2 - 2\langle u, v\rangle + \|v\|^2) + p^2 \\ &= 2\langle u, v\rangle - \|v\|^2 + p^2. \end{aligned}$$

Luego, tenemos que

$$\begin{aligned} 2\langle u, v\rangle &= \|v\|^2 + r^2 - p^2 \\ &= \|v\|^2 + (r + p)(r - p) \\ &= \|v\|(\|v\| + r - p) \\ &= \|v\|(r + p + r - p) = 2\|v\|r = 2\|v\|\|u\| \cos(\theta). \end{aligned}$$

Dividiendo por dos a ambos lados, se deduce el resultado.

Caso 5, $\theta \in (\frac{\pi}{2}, \pi)$**:** Finalmente, analizamos el caso en que el ángulo θ es obtuso, es decir, $\theta \in (\pi/2, \pi)$. Recordando que el ángulo θ es el que se forma entre u y la recta generada por v, y definiendo θ' al ángulo entre u y $-v$, se tiene que $\theta + \theta' = \pi$, como se ilustra en la Figura 1.7. Como θ es obtuso,

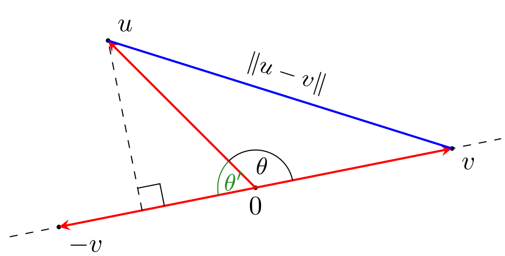


Figura 1.7: El caso donde θ es obtuso, se puede reemplazar v por $-v$ y θ por $\theta' = \pi - \theta$ para reducirse al caso agudo.

$\theta' = \pi - \theta$ es agudo. Esto nos permite aplicar el desarrollo anterior reemplazando v por $-v$, concluyendo que

$$\langle u, -v\rangle = \|u\|\| - v\| \cos(\theta') = \|u\|\|v\| \cos(\pi - \theta)$$

Recordando que $\cos(\pi - \theta) = -\cos(\theta)$ podemos escribir, gracias a la bilinealidad del producto escalar (Proposición 1.1.8), que

$$-\langle u, v\rangle = \langle u, -v\rangle = \|u\|\|v\| \cos(\pi - \theta) = -\|u\|\|v\| \cos(\theta).$$

Multiplicando a amobs lados por -1, se concluye el resultado. Esto termina la demostración. $\square$

Atención

Algunos libros definen el ángulo entre dos vectores u y v de $\mathbb{R}^n$ (distintos de cero) como

$$\theta = \arccos\left(\frac{\langle u, v\rangle}{\|u\|\|v\|}\right),$$

tomando el Teorema 1.1.11 como una definición. En este apunte preferimos seguir un enfoque más geométrico, tomando la Definición 1.1.10 que viene dada de la geometría euclidiana en $\mathbb{R}^2$.

Del teorema 1.1.11 se desprenden dos importantes consecuencias, muy útiles en el estudio topológico de $\mathbb{R}^n$: 1) La desigualdad de Cauchy-Schwarz, que nos permite acotar el producto escalar de vectores por el producto de sus normas; y 2) el Teorema del coseno, que generaliza el teorema de Pitágoras a triángulos cualquiera (formados por el origen y dos vectores).

Corolario 1.1.12 (Desigualdad de Cauchy-Schwartz)**.** *Sea $n \in \mathbb{N}$, y sean u y v dos vectores en $\mathbb{R}^n$. Se tiene que*

$$|\langle u, v\rangle| \leq \|u\|\|v\|. \tag{1.15}$$

Demostración. En el caso que $n = 1$, u y v serían números escalares, por lo que podemos escribir directamente

$$|\langle u, v\rangle| = |uv| = |u||v| = \|u\|\|v\|.$$

Para el caso de $n \in \mathbb{N}$ general, si alguno de los dos vectores es cero, la desigualdad es trivial. Si ambos son distintos de cero, Podemos aplicar el Teorema 1.1.11, escribiendo

$$|\langle u, v\rangle| = |\|u\|\|v\| \cos(\theta)| = \|u\|\|v\||\cos(\theta)| \leq \|u\|\|v\|,$$

donde la última desigualdad se obtiene del hecho que $0 \leq |\cos(\theta)| \leq 1$. Esto concluye la demostración. □

Corolario 1.1.13 (Teorema del coseno)**.** *Sea $n \in \mathbb{N}$, y sean u y v dos vectores en $\mathbb{R}^n$ distintos de cero. Sea θ el ángulo entre u y v. Se tiene que*

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos(\theta). \tag{1.16}$$

Demostración. Basta con ocupar la Ley del paralelogramo (Proposición 1.1.9) y el Teorema 1.1.11, escribiendo

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos(\theta).$$

□

1.2 Sucesiones

Hasta ahora, hemos visto que el cálculo en $\mathbb{R}$ involucra variables continuas. Por ejemplo, al ver un polinomio cuadrático de forma $p(x) = ax^2 + bx + c$, entendemos que la variable x puede tomar valores reales; o al ver una función trigonométrica como $\cos(\theta)$, interpretamos normalmente que el ángulo θ puede tomar valores en un intervalo real, como $[0, 2\pi)$. Estas son variables continuas.

Sin embargo, en muchas situaciones, nos encontramos con la necesidad de considerar funciones que tomen valores *discretos*. En algoritmos computacionales, normalmente generamos secuencias de números que se van aproximando a la solución de un problema. Algunos ejemplos de esta idea consiste en cálculos numéricos de integrales, algoritmos de optimización basados en métodos de descenso, o el aprendizaje de máquinas. También en mediciones experimentales, registramos una secuencia de mediciones en instantes discretos pues somos incapaces de medir de manera continua un fenómeno. En general, los datos de los que disponemos cuando hacemos análisis de datos tienen forma de secuencia de números, sin tener una fórmula analítica que gobierne su comportamiento.

Para poder hacer cálculo con variables discretas, se introduce la noción de *sucesión*. En esta sección, estudiaremos rápidamente cómo funcionan los límites de sucesiones en una dimensión, y cómo se relacionan con los límites de funciones a variable real.

1.2.1 Conceptos y Propiedades

Definición 1.2.1 (Sucesión real). *Una sucesión real es una función que va de $\mathbb{N}$ a $\mathbb{R}$, es decir,*

$$\begin{aligned} x : \mathbb{N} &\to \mathbb{R} \\ k &\mapsto x(k) \in \mathbb{R}, \end{aligned}$$

donde denotamos cada elemento con un subíndice que nos indique su orden $x_k := x(k)$ y la sucesión completa por $(x_k)_{k\in\mathbb{N}}$, o $(x_k)_k$, o incluso (x_k), cuando no haya posibilidad de confusión.

Atención

- En general, para las sucesiones se usan las letras s, u, v, x, y, z.
- En general, para los subíndices se usan i, j, k, n, m.
- En algunos casos, las sucesiones están bien definidas a partir de cierto número. Por ejemplo, la sucesión $x_k = \arccos(20/k)$ está bien definida a partir de $\bar{k} = 20$. En estos escribiremos $(x_k)_{k\geq\bar{k}}$ para explicitar que la sucesión toma índices $k \in \mathbb{N}$ a partir del primer índice es $\bar{k} \in \mathbb{N}$.
- Dada una sucesión $(x_k)_{k\geq\bar{k}}$, es posible definir otra sucesión en todos los naturales $(y_k)_{k\in\mathbb{N}}$ que represente la misma secuencia de números, con el cambio de variables dado por

$$y_k = x_{k+\bar{k}-1}.$$

Ejemplo 1.2.2. Algunos ejemplos de sucesiones son

1. $(x_k)_{k\in\mathbb{N}}$ dada por $x_k = 1/k$, es la sucesión $(x_k) = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4} \ldots\right)$.
2. $(x_k)_{k\in\mathbb{N}}$ dada por $x_k = \frac{k^2+\ln(k)}{\sqrt{k}}$ es la sucesión dada por $(1, 3.3186, 5.8304, 8.6931, \ldots)$.
3. $(y_k)_{k\in\mathbb{N}}$ dada por $y_k = (-1)^k$ es la sucesión dada por $(-1, 1, -1, 1, \ldots)$.
4. $(z_i)_{i\geq 3}$ dada por $z_i = (-1)^i(i-3)$ es la sucesión dada por $(0, 1, -2, 3, \ldots)$.
5. $s_n = \sqrt{(-1)^n}$ representa una función que se indefine término por medio, ya que $(s_n) = (1, \nexists, 1, \nexists, 1, \nexists, \ldots)$. Es decir, se indefine una cantidad infinita de términos por lo que **no es una sucesión**.

Notar que los cambios de letra de la sucesión y de letra del índice son solo notacionales. Incluso el punto de partida de la sucesión puede entenderse como notacional, considerando el cambio de variable descrito en el recuadro de Atención anterior. ♢

Definición 1.2.3 (Convergencia de sucesiones reales). *Decimos que una sucesión $(x_k)_k$ en $\mathbb{R}$ **converge** a un real $\bar{x} \in \mathbb{R}$, si*

$$(\forall\varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \quad |x_k - \bar{x}| \leq \varepsilon. \tag{1.17}$$

*En caso de que lo anterior se cumpla, $\bar{x}$ se denomina **límite** de $(x_k)_k$ y escribimos*

$$\lim_{k\to\infty} x_k = \bar{x} \quad y \quad x_k \xrightarrow{k\to\infty} \bar{x}.$$

La ecuación (1.17) puede leerse de la siguiente manera: Para cualquier tolerancia $\varepsilon > 0$, existe un momento $k_0 \in \mathbb{N}$ tal que, a partir de ese momento en adelante ($k \geq k_0$), todos los elementos de la sucesión están en el intervalo $[\bar{x} - \varepsilon, \bar{x} + \varepsilon]$ (están a lo más a distancia ε de $\bar{x}$). Dicho de otro modo, la sucesión $(x_k)_k$ converge a $\bar{x}$ si en la medida que "el índice k avanza", la sucesión logra estar "arbitrariamente cerca" de $\bar{x}$.

Nota: Pasos de demostración de convergencia

Supongamos que tenemos una sucesión $(x_k) \subseteq \mathbb{R}$ y queremos demostrar que $x_k \to \bar{x}$ usando la Definición 1.2.3. Para esto debemos seguir los siguientes pasos:

> **Paso 1:** Fijar $\varepsilon > 0$ arbitrario.
>
> **Paso 2:** Construir $k_0 \in \mathbb{N}$, tal que cada vez que tomamos un $k \geq k_0$, se cumple que $|x_k - \bar{x}| \leq \varepsilon$. **La construcción de k_0 dependerá de ε y de de las características de la sucesión misma.**
>
> **Paso 3:** Notar que como $\varepsilon > 0$ es arbitrario, la construcción anterior muestra lo siguiente: Para cualquier $\varepsilon > 0$, somos capaces de construir un $k_0 \in \mathbb{N}$ que cumple que para todo $k \geq k_0$, se verifica la desigualdad $|x_k - \bar{x}| \leq \varepsilon$. Esto es exactamente lo que está escrito en (1.17) para la sucesión (x_k) y el límite $\bar{x}$.
>
> **Paso 4:** Notar que como demostramos (1.17), esto es lo mismo que $x_k \to \bar{x}$.

En las demostraciones de esta sección, ocuparemos esta estrategia recurrentemente. En esta estrategia, el núcleo de la demostración es el paso 2.

Esta definición de límites de hecho es análoga a la definición de límites de funciones $f(t)$ cuando $t \to +\infty$, con la diferencia que f sólo está siendo evaluada en los puntos $t = k \in \mathbb{N}$. Esta relación se captura en la siguiente proposición.

Proposición 1.2.4. *Sea $f : \mathbb{R} \to \mathbb{R}$ una función y $\alpha \in \mathbb{R}$ tales que*

$$\lim_{t \to +\infty} f(t) = \alpha.$$

Entonces, definiendo la sucesión $(x_k)_k$ dada por $x_k = f(k)$, se tiene que $x_k \xrightarrow{k \to \infty} \alpha$.

Demostración. Recordemos que la definición de límite en el infinito de una función real es la siguiente:

$$\lim_{t \to +\infty} f(t) = \alpha \iff (\forall \varepsilon > 0)\, (\exists M > 0)\, (\forall t \geq M)\ |f(t) - \alpha| \leq \varepsilon. \tag{1.18}$$

Nosotros queremos demostrar que $f(k) \to \alpha$, o equivalentemente, que la ecuación (1.17) se verifica para $x_k = f(k)$ y $\bar{x} = \alpha$. Es decir, queremos demostrar que:

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)\ |f(k) - \alpha| \leq \varepsilon.$$

<u>Tomemos $\varepsilon > 0$ arbitrario</u>. Usando (1.18), sabemos que existe $M > 0$ (que depende de ε) tal que

$$\forall t \geq M,\ |f(t) - \alpha| \leq \varepsilon.$$

<u>Definamos $k_0 = \lceil M \rceil$</u>, donde $\lceil M \rceil$ es el cajón superior de M. Para todo $k \geq k_0$, tenemos que tomando $t = k$ en la desigualdad anterior, podemos escribir

$$k \geq k_0 \implies k \geq M \implies |f(k) - \alpha| \leq \varepsilon.$$

Luego, se cumple que

$$\forall k \geq k_0, |f(k) - \alpha| \leq \varepsilon.$$

Notando que esta construcción considera $\varepsilon > 0$ arbitrario, concluimos lo siguiente: Para todo $\varepsilon > 0$, existe $k_0 \in \mathbb{N}$ (el que se obtiene siguiendo el desarrollo anterior), tal que para todo $k \geq k_0$ se cumple que $|f(k) - \alpha| \leq \varepsilon$. Es decir, (1.17) se verifica para $x_k = f(k)$ y $\bar{x} = \alpha$, que es exactamente lo que queríamos demostrar. □

Con esta idea en mente, podemos replicar para límites de sucesiones varios de los resultados que tenemos para límites de funciones, tales como Unicidad del límite, Álgebra de límites, y Teorema del Sándwich.

Proposición 1.2.5 (Unicidad del límite de sucesiones reales)**.** *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$. Si el límite de $(x_k)_{k\in\mathbb{N}}$ existe, entonces es único.*

Demostración. Razonemos por contradicción, y supongamos que la sucesión $(x_k)_{k\in\mathbb{N}}$ tiene dos límites distintos, $\bar{x}_1$ y $\bar{x}_2$. Sea $\varepsilon > 0$ arbitrario.

- Como $x_k \to \bar{x}_1$, usando (1.17) con $\varepsilon/2$, existe $k_1 \in \mathbb{N}$ tal que $\forall k \geq k_1 : \ |x_k - \bar{x}_1| \leq \varepsilon/2$.
- Como $x_k \to \bar{x}_2$, usando (1.17) con $\varepsilon/2$, existe $k_2 \in \mathbb{N}$ tal que $\forall k \geq k_2 : \ |x_k - \bar{x}_2| \leq \varepsilon/2$.

Tomemos $k_0 = \text{máx}\{k_1, k_2\}$. Como las dos condiciones de arriba se cumplen para $k = k_0$, podemos usar la desigualdad triangular de la distancia (ver Proposición 1.1.6) para escribir

$$d(\bar{x}_1, \bar{x}_2) \leq d(\bar{x}_1, x_{k_0}) + d(x_{k_0}, \bar{x}_2) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Luego, como $\varepsilon > 0$ es arbitrario, podemos concluir que

$$d(\bar{x}_1, \bar{x}_2) \leq \varepsilon, \ \forall \varepsilon > 0.$$

Esto implica que necesariamente $d(\bar{x}_1, \bar{x}_2) = 0$. Por lo tanto, usando la propiedad de separación de la distancia (ver Proposición 1.1.6), concluimos que $\bar{x}_1 = \bar{x}_2$, lo cual es una contradicción. Esto concluye la demostración. □

Definición 1.2.6 (Sucesión Nula)**.** *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$. La llamaremos nula si (x_k) converge a cero, es decir,*

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \quad |x_k| \leq \varepsilon$$

Definición 1.2.7 (Sucesión Acotada)**.** *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$. La llamaremos acotada si:*

$$(\exists M > 0)(\forall n \in \mathbb{N}) \quad |x_n| \leq M$$

Proposición 1.2.8. *Sean (x_k) e (y_k) sucesiones en $\mathbb{R}$. Entonces,*

1. *(x_k) es nula si, y sólo si, $(|x_k|)$ es nula.*
2. *Si (x_k) es nula, entonces (x_k) es acotada.*
3. *Si (x_k) e (y_k) son nulas, entonces la sucesión (z_k) dada por $z_k = x_k + y_k$, también es nula.*
4. *Si (x_k) e (y_k) son acotadas, entonces la sucesión (z_k) dada por $z_k = x_k + y_k$, también es acotada.*
5. *Si (x_k) e (y_k) son nulas, entonces la sucesión (z_k) dada por $z_k = x_k \cdot y_k$, también es nula.*
6. *Si (x_k) e (y_k) son acotadas, entonces la sucesión (z_k) dada por $z_k = x_k \cdot y_k$, también es acotada.*
7. *Si (x_k) es nula e (y_k) es acotada, entonces la sucesión (z_k) dada por $z_k = x_k \cdot y_k$ es nula.*

Demostración. Vamos a demostrar cada afirmación por separado.

Propiedad 1: Escribamos que $(|x_k|)$ es nula según la Definición 1.2.6:

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0) \quad |(|x_k|)| \leq \varepsilon.$$

Como $|(|x_k|)| = |x_k|$, podemos ver que obtenemos defición de que (x_k) es nula, por lo que son equivalentes.

Propiedad 2: Queremos demostrar que (x_k) es acotada. Como (x_k) es nula, tomando $\varepsilon = 1$ en la Definición 1.2.6, sabemos que

$$(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)\ |x_k| \leq 1$$

Tomemos el máximo de los valores anteriores a k_0 de la sucesión (en valor absoluto), es decir, $m = \text{máx}\{|x_k| : \ k < k_0\}$. Por lo que tenemos que la primera parte de la sucesión (hasta k_0) está acotada por m, y la segunda parte de la sucesión (a partir de k_0) está acotada por 1. Tomando $M = \text{máx}\{m, 1\}$, podemos concluir que:

$$\forall k \in \mathbb{N} : \quad |x_k| \leq M$$

que es lo que queríamos demostrar.

Propiedad 3: Queremos demostrar que $(z_k) = (x_k + y_k) \to 0$. Sea $\varepsilon > 0$ arbitrario. Como (x_k) e (y_k) son nulas, podemos ocupar la Definición 1.2.6 con $\varepsilon/2$, y obtenemos:

- $(\exists k_1 \in \mathbb{N})(\forall k \geq k_1)\ |x_k| \leq \varepsilon/2$
- $(\exists k_2 \in \mathbb{N})(\forall k \geq k_2)\ |y_k| \leq \varepsilon/2$

Tomando $k_0 = \text{máx}\{k_1, k_2\}$ y utilizando desigualdad triangular, se tiene que para todo $k \geq k_0$:

$$|z_k| = |x_k + y_k| \leq |x_k| + |y_k| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Luego, como ε es arbitrario podemos concluir que

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)|z_k| \leq \varepsilon$$

Es decir que $z_k \to 0$.

Propiedad 4: Queremos demostrar que $(z_k) = (x_k + y_k)$ es acotada. Como (x_k) e (y_k) son acotadas tenemos que:

- $(\exists M_1 > 0)(\forall k \in \mathbb{N})\ |x_k| \leq M_1$
- $(\exists M_2 > 0)(\forall k \in \mathbb{N})\ |y_k| \leq M_2$

Tomando $M = M_1 + M_2$ y utilizando desigualdad triangular, se tiene que para todo $k \in \mathbb{N}$:

$$|z_k| = |x_k + y_k| \leq |x_k| + |y_k| \leq M_1 + M_2 = M,$$

con lo que concluimos que (z_k) es acotada.

Propiedad 5: La demostración es análoga a la Propiedad 3 tomando $\sqrt{\varepsilon}$ para cada sucesión.

Propiedad 6: La demostración es análoga a la Propiedad 4 tomando $M = M_1 \cdot M_2$.

Propiedad 7: Queremos demostrar que $(z_k) = (x_k \cdot y_k) \to 0$. Sea $\varepsilon > 0$ arbitrario.
Como (y_k) es acotada, exite $M > 0$ tal que $|y_k| \leq M$ para todo $k \in \mathbb{N}$.
Como (x_k) es nula, utilizando ε/M en la Definición 1.2.6 tenemos que:

$$(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)\ |x_k| \leq \frac{\varepsilon}{M}$$

Tomando el mismo k_0 se tiene que para todo $k \geq k_0$:

$$|z_k| = |x_k \cdot y_k| = |x_k| \cdot |y_k| \leq \frac{\varepsilon}{M} \cdot M = \varepsilon,$$

Luego, como ε es arbitrario podemos concluir que

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)|z_k| \leq \varepsilon$$

Es decir que $(z_k) \to 0$.

□

Proposición 1.2.9. *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$. Entonces, $x_k \to \bar{x} \iff (z_k)_{k\in\mathbb{N}}$, donde $z_k = x_k - \bar{x}$, es una sucesión nula.*

Demostración. Escribamos, usando (1.17) para $x_k \to \bar{x}$ tenemos:

$$(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)|x_k - \bar{x}| \leq \varepsilon$$

que es equivalente a que $z_k \to 0$. □

Proposición 1.2.10. *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$ tal que $x_k \to \bar{x}$. Entonces, (x_k) es una sucesión acotada.*

Demostración. Sea $(x_k)_{k\in\mathbb{N}}$ tal que $x_k \to \bar{x}$. Por la Proposición 1.2.9, tenemos que $(x_k - \bar{x})$ es nula. Utilizando la propiedad 2 de la Proposición 1.2.8 tenemos que $(x_k - \bar{x})$ es acotada, es decir:

$$(\exists M' > 0)(\forall k \in \mathbb{N})\ |x_k - \bar{x}| \leq M$$

Tomando $M = M' + |\bar{x}|$, y utilizando la desigualdad triángular tenemos que para todo $k \in \mathbb{N}$:

$$|x_k| = |x_k - \bar{x} + \bar{x}| \leq |x_k - \bar{x}| + |\bar{x}| \leq M' + \bar{x} = M,$$

que es lo que necesitamos para concluir que la sucesión (x_k) es acotada.

□

La Proposición 1.2.8, en particular la Propiedad 7, es muy útil para concluir que alteraciones a sucesiones nulas conocidad siguen siendo nulas. Aquí será importante la Proposición 1.2.10 para poder usar la propiedad 7. Finalmente, con la Propiedad 1.2.9 vamos a poder concluir convergencia de sucesiones sin tener que pasar por (1.17).

Proposición 1.2.11 (Álgebra de sucesiones reales)**.** *Sean $(x_k)_{k\in\mathbb{N}}, (y_k)_{k\in\mathbb{N}}$ dos sucesiones en $\mathbb{R}$ tal que $\lim x_k = \bar{x}$ y $\lim y_k = \bar{y}$. Se tienen las siguientes propiedades:*

1. *La sucesión (z_k) dada por $z_k = x_k + y_k$ cumple que $\lim_k z_k = \bar{x} + \bar{y}$.*
2. *La sucesión (z_k) dada por $z_k = x_k \cdot y_k$ cumple que $\lim_k z_k = \bar{x} \cdot \bar{y}$.*
3. *Si $\bar{y} \neq 0$, entonces existe $k_0 \in \mathbb{N}$ tal que la sucesión $z_k = \frac{x_k}{y_k}$ está bien definida para todo $k \geq k_0$. Más aún, la sucesión $(z_k)_{k\geq k_0}$ converge a $\bar{z} = \frac{\bar{x}}{\bar{y}}$.*

Demostración. Vamos a demostrar cada afirmación por separado.

Propiedad 1: Queremos demostrar que $z_k \to \bar{x} + \bar{y}$. Combinando las proposiciones anteriores podemos realizar el siguiente razonamiento:

- Por Proposición 1.2.9: como $x_k \to \bar{x}$, entonces $(x_k - \bar{x})$ es nula.
- Por Proposición 1.2.9: como $y_k \to \bar{y}$, entonces $(y_k - \bar{y})$ es nula.
- Por Proposición 1.2.8 (Prop. 3): $((x_k - \bar{x}) + (y_k - \bar{y}))$ es nula.
- Reordenando términos: $(x_k + y_k - (\bar{x} + \bar{y}))$ es nula.
- Identificando a z_k : $(z_k - (\bar{x} + \bar{y}))$ es nula.
- Por Proposición 1.2.9 concluimos que $z_k \to (\bar{x} + \bar{y})$.

Propiedad 2: Para esta demostración vamos a realizar la **misma demostración** en sentido directo e inverso, pues para algunos lectores puede ser más fácil de entender una sobre otra.

Demostración inversa Queremos demostrar que $z_k \to \bar{x} \cdot \bar{y}$.

- Es equivalente a que $(z_k - \bar{x} \cdot \bar{y})$ sea nula.
- Usando la definición de z_k, es equivalente a que $(x_k \cdot y_k - \bar{x} \cdot \bar{y})$ sea nula.
- Sumando y restando $\bar{x} \cdot y_k$, es equivalente a que $(x_k \cdot y_k - \bar{x} \cdot y_k + \bar{x} \cdot y_k - \bar{x} \cdot \bar{y})$ sea nula.
- Factorizando tenemos:
$$\underbrace{(x_k - \bar{x}) \cdot y_k}_{(1)} + \underbrace{(y_k - \bar{y}) \cdot \bar{x}}_{(2)} \tag{1.19}$$

  (1) Es nula porque
    - $x_k \to \bar{x} \implies (x_k - \bar{x})$ es nula.
    - (y_k) converge $\implies (y_k)$ es acotada
    - se concluye por *sucesión nula por sucesión acotada.*

  (2) Es nula porque
    - $y_k \to \bar{y} \implies (y_k - \bar{y})$ es nula.
    - Una constante (en este caso $\bar{x}$) se puede considerar como una sucesión constante, que siempre es acotada.
    - se concluye por *sucesión nula por sucesión acotada.*

  Por suma de sucesiones nulas, tenemos que (1.19) es efectivamente nula, y podemos concluir que $z_k \to \bar{x} \cdot \bar{y}$

Demostración directa Queremos demostrar que $z_k \to \bar{x} \cdot \bar{y}$

- Como $x_k \to \bar{x}$, entonces $(x_k - \bar{x})$ es nula.
- Como (y_k) es convergente, entonces también es acotada.
- $(x_k - \bar{x}) \cdot y_k$ es *sucesión nula por sucesión acotada*, por lo que es nula. (1)
- Como $y_k \to \bar{y}$, entonces $(y_k - \bar{y})$ es nula.
- Una constante es, por definición, acotada. Por lo que $(y_k - \bar{y}) \cdot \bar{x}$ se puede considerar una sucesión nula por una acotada, por lo que la sucesión resultante es nula. (2)

- Sumando (1) y (2), por suma de sucesiones nulas, tenemos que la siguiente sucesión es nula:

$$((x_k - \bar{x}) \cdot y_k) + ((y_k - \bar{y}) \cdot \bar{x})$$

- Distribuyendo se obtiene la sucesión $(x_k \cdot y_k - \bar{x} \cdot y_k + \bar{x} \cdot y_k - \bar{x} \cdot \bar{y})$
- Reordenando tenemos que $(x_k \cdot y_k - \bar{x} \cdot \bar{y})$ es nula.
- Identificando a z_k tenemos que $(z_k - \bar{x} \cdot \bar{y})$ es nula.
- Podemos concluir que $z_k \to \bar{x} \cdot \bar{y}$.

Propiedad 3: Supongamos que $\bar{y} \neq 0$. Para ver que está bien definida queremos ver que la sucesión (y_k) es distinta de cero a partir de cierto punto.

Para cualquier término de la sucesión:

$$|\bar{y}| - |y_k| = |\bar{y} - y_k + y_k| - |y_k| \leq |\bar{y} - y_k| + |y_k| - |y_k| = |\bar{y} - y_k|.$$

Como $y_k \to \bar{y}$, podemos usar (1.17) con $\varepsilon = |\bar{y}|/2$, por lo que existe $k_0 \in \mathbb{N}$ tal que

$$\forall k \geq k_0,\ \ |\bar{y} - y_k| \leq \varepsilon = \frac{|\bar{y}|}{2}.$$

Uniendo ambas desigualdades anteriores obtenemos:

$$\begin{aligned} \forall k \geq k_0,\ |\bar{y}| - |y_k| &\leq |\bar{y} - y_k| \leq \frac{|\bar{y}|}{2} \\ \implies |\bar{y}| - \frac{|\bar{y}|}{2} &\leq |y_k| \\ \implies \frac{|\bar{y}|}{2} &\leq |y_k|. \end{aligned}$$

Por lo tanto concluimos que:

$$\forall k \geq k_0,\ 0 < \frac{|\bar{y}|}{2} \leq |y_k|. \tag{1.20}$$

Como (y_k) es distinta de cero a partir de k_0, entonces se puede dividir por $(y_k)_{k \geq k_0}$. Concluimos $(z_k)_{k \geq k_0}$ está bien definida.

Veamos ahora la convergencia, vamos a razonar de nuevo de manera inversa:

- Queremos que $z_k \to \frac{\bar{x}}{\bar{y}}$.
- Es equivalente a que $\left(z_k - \frac{\bar{x}}{\bar{y}}\right)$ sea nula.
- Es equivalente a que $\left(\frac{x_k}{y_k} - \frac{\bar{x}}{\bar{y}}\right)$ sea nula.
- Manipulando la sucesión obtenemos:

$$\begin{aligned} \frac{x_k}{y_k} - \frac{\bar{x}}{\bar{y}} &= \frac{x_k\bar{y} - y_k\bar{x}}{y_k \cdot \bar{y}} \\ &= \frac{x_k\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} - y_k\bar{x}}{y_k \cdot \bar{y}} \\ &= \underbrace{\frac{1}{\bar{y}} \cdot \frac{1}{y_k}}_{(1)} \cdot \underbrace{(\bar{y}(x_k - \bar{x}) - \bar{x}(y_k - \bar{y}))}_{(2)} \end{aligned}$$

(1) Es la constante $\frac{1}{\bar{y}}$ (acotada) por la sucesión $\left(\frac{1}{y_k}\right)_{k\geq k_0}$. Tomando inversa de la desigualdad (1.20) tenemos que $\left|\frac{1}{y_k}\right| \leq \frac{2}{|\bar{y}|} = M$, por lo que $\left(\frac{1}{y_k}\right)_{k\geq k_0}$ es una sucesión acotada. Por *multiplicación de sucesiones acotadas*, podemos concluir que (1) es acotada.

(2) Es una sucesión nula por el mismo argumento de (1.19).

Por lo tanto, utilizando *sucesión nula por sucesión acotada*, obtenemos que la sucesión es nula.

- Podemos concluir que $z_k \to \frac{\bar{x}}{\bar{y}}$.

□

Otro elemento clave de la manipulación y transformación de límites de sucesiones tiene que ver con las funciones que podemos "*aplicar*" a las sucesiones.

Proposición 1.2.12. *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}$ tal que $x_k \to \bar{x}$, y $f : \mathbb{R} \to \mathbb{R}$ es una función continua en $\bar{x}$. Entonces la sucesión (z_k) dada por $z_k = f(x_k)$ converge a $f(\bar{x})$.*

Demostración. Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua en $\bar{x}$. Esto quiere decir que $\lim_{x\to\bar{x}} f(x) = f(\bar{x})$. Ocupando la caracterización ε-δ de límite de funciones, tenemos que el límite anterior es equivalente a escribir

$$\forall\varepsilon > 0, \exists\delta > 0, \big(|x - \bar{x}| \leq \delta \implies |f(x) - f(\bar{x})| \leq \varepsilon\big). \tag{1.21}$$

Sea ahora $\varepsilon > 0$ arbitrario, y sea $\delta > 0$ aquel que verifica (1.21). Como $x_k \to \bar{x}$, entonces existe $k_0 \in \mathbb{N}$ tal que

$$\forall k \geq k_0,\ |x_k - \bar{x}| \leq \delta.$$

Mezclando la desigualdad anterior con (1.21), podemos escribir que

$$\forall k \geq k_0,\ |f(x_k) - f(\bar{k})| \leq \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, concluimos que

$$\forall\varepsilon > 0, \exists k_0 \in \mathbb{N}\forall k \geq k_0,\ |f(x_k) - f(\bar{k})| \leq \varepsilon.$$

Esto equivale a decir que $f(x_k) \to f(\bar{x})$. Esto concluye la demostración. □

Para finalizar esta sección, presentaremos el Teorema del Sándwich, que es uno de los teoremas más útiles para demostrar convergencia de sucesiones. Este teorema establece que si somos capaces de acotar una sucesión (y_k) por arriba y por abajo por sucesiones (x_k) y (z_k) con un mismo límite, entonces ese límite común también es el límite de (y_k).

Proposición 1.2.13 (Teorema del Sándwich para sucesiones reales)**.** *Sean $(x_k)_{k\in\mathbb{N}}, (y_k)_{k\in\mathbb{N}}$ y $(z_k)_{k\in\mathbb{N}}$ tres sucesiones en $\mathbb{R}$ tales que*

(i) Existe $k_0 \in \mathbb{N}$ tal que $x_k \leq y_k \leq z_k$ para todo $k \geq k_0$

(ii) $\lim_k x_k = \lim_k z_k = \alpha \in \mathbb{R}$.

Se tiene entonces que la sucesión $(y_k)_k$ es convergente con $\lim_k y_k = \alpha$.

Demostración. <u>Sea $\varepsilon > 0$ arbitrario</u>. Como $x_k \to \alpha$ y $z_k \to \alpha$, existen $k_1, k_2 \in \mathbb{N}$ tales que

$$\forall k \geq k_1,\ |x_k - \alpha| \leq \varepsilon \quad \text{y} \quad \forall k \geq k_2,\ |z_k - \alpha| \leq \varepsilon.$$

<u>Tomemos $k_3 = \max\{k_0, k_1, k_2\}$.</u> De las desigualdades anteriores tenemos que

$$\forall k \geq k_3, \quad -\varepsilon \leq x_k - \alpha \ \text{ y } \ z_k - \alpha \leq \varepsilon.$$

Luego, ocupando que $x_k \leq y_k \leq z_k$ para todo $k \geq k_0$, podemos escribir que

$$\forall k \geq k_3, \quad -\varepsilon \leq x_k - \alpha \leq y_k - \alpha \leq z_k - \alpha \leq \varepsilon.$$

Por lo que $|y_k - \alpha| \leq \varepsilon$ para todo $k \geq k_3$. Como $\varepsilon > 0$ es arbitrario, esto muestra que $y_k \to \alpha$, lo que concluye la demostración. $\square$

Nota: Rango de sucesiones

Cuando una sucesión $(x_k)_{k\in\mathbb{N}}$ toma todos sus valores en un conjunto de A de $\mathbb{R}$, podemos escribir $(x_k)_{k\in\mathbb{N}} \subseteq A$, para enfatizar esta inclusión. Es decir,

$$(x_k)_{k\in\mathbb{N}} \subseteq A \iff \forall k \in \mathbb{N},\ x_k \in A. \tag{1.22}$$

1.3 Sucesiones en $\mathbb{R}^n$

En dimensiones superiores, también podemos definir lo que sería una sucesión, de la misma manera que lo hicimos en el caso de sucesiones reales.

1.3.1 Conceptos y Propiedades

Definición 1.3.1 (Sucesión en $\mathbb{R}^n$)**.** *Una sucesión en $\mathbb{R}^n$ (o sucesión vectorial) es una función*

$$\begin{aligned} x : \mathbb{N} &\to \mathbb{R}^n \\ k &\mapsto x(k), \end{aligned}$$

*donde denotamos a cada **vector** de la sucesión con subíndices según su orden $x_k := x(k)$. Para la sucesión completa, también ocuparemos las notaciones $(x_k)_{k\in\mathbb{N}}$, o $(x_k)_k$ o incluso (x_k), cuando no haya posibilidad de confusión.*

Nota: Índices en sucesiones vectoriales

Al tener una sucesión (x_k) en $\mathbb{R}^n$, estamos diciendo que cada elemento x_k de la sucesión es un vector en $\mathbb{R}^n$. Es decir, x_k tiene n coordenadas. Para poder hablar de las coordenadas de los elementos de la sucesión (x_k), necesitamos ocupar un segundo índice.

En general ocuparemos la siguiente notación para sucesiones de vectores:

- w, x, y, z para las sucesiones.
- n, m, d para el índices de que refiere a la dimensión ($\mathbb{R}^n, \mathbb{R}^m$, etc.).
- i, j, k para los indices que recorren tanto la sucesión, como cada vector.

Por ejemplo $(x_k) \subseteq \mathbb{R}^n$ es una sucesión de vectores en $\mathbb{R}^n$. El índice de la sucesión es k que recorre todos los naturales $(1, 2, 3, ...)$. Para describir cada vector necesitamos un nuevo índice, elijamos i, que recorre $[n]$. Entonces el primer término, y el k-ésimo término de la sucesión son:

$$x_1 = \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{1,i} \\ \vdots \\ x_{1,n} \end{pmatrix} = (x_{1,i})_{i=1}^n \in \mathbb{R}^n, \qquad x_k = \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,i} \\ \vdots \\ x_{k,n} \end{pmatrix} = (x_{k,i})_{i=1}^n \in \mathbb{R}^n. \tag{1.23}$$

Recordemos que las sucesiones en $\mathbb{R}$ las podemos interpretar como una secuencia ordenada de puntos en la recta real. Similarmente, una sucesión en $\mathbb{R}^n$ se puede interpretar como una secuencia ordenada de puntos en el espacio $\mathbb{R}^n$. Por supuesto, el orden de la secuencia no está asociado a un orden en la recta o en el espacio $\mathbb{R}^n$, si no que se refiere a que los puntos están indexados por los números naturales, lo cual les da un orden: Si $k_1 < k_2$ entonces el punto x_{k_1} viene primero en la sucesión, y luego viene el punto x_{k_2}.

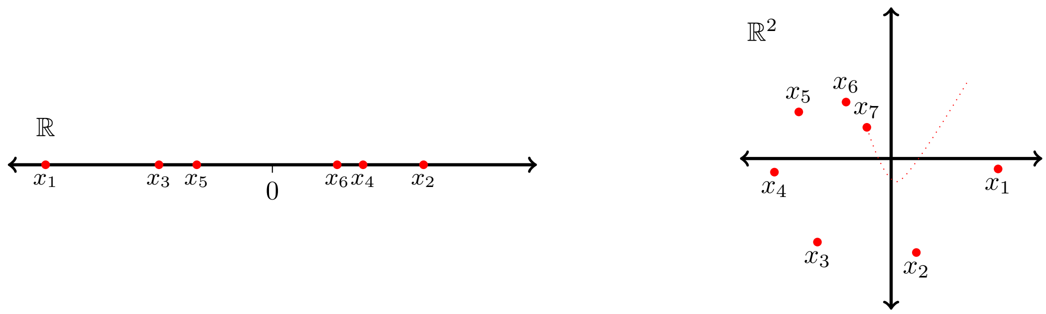


Figura 1.8: Esquema de sucesiones en $\mathbb{R}$ y en $\mathbb{R}^2$.

Otra perspectiva que también ocuparemos en este apunte es pensar que las sucesiones vectoriales se componen de varias sucesiones reales: Para una sucesión $(x_k)_k \subseteq \mathbb{R}^n$, cada coordenada $i \in [n]$ induce una sucesión real $(y_k) \subseteq \mathbb{R}$ dada por $y_k = x_{k,i}$. Esquemáticamente, podemos ver esta doble interpretación en el siguiente esquema:

$$x_k = \begin{pmatrix} x_{k,1} \\ \vdots \\ x_{k,i} \\ \vdots \\ x_{k,n} \end{pmatrix} \begin{array}{l} \xrightarrow{\text{sucesión primera coordenada}} (x_{k,1})_{k\in\mathbb{N}} \\ \qquad\qquad\vdots \\ \xrightarrow{\text{sucesión i-ésima coordenada}} (x_{k,i})_{k\in\mathbb{N}} \\ \qquad\qquad\vdots \\ \xrightarrow{\text{sucesión n-ésima coordenada}} (x_{k,n})_{k\in\mathbb{N}} \end{array} \tag{1.24}$$

Ejemplo 1.3.2. Algunos ejemplos de sucesiones vectoriales son:

1. La sucesión $(x_k) \subseteq \mathbb{R}^2$ dada por
$$x_k = \left(1-k, k^2\right).$$
Sus elementos son:
$$x_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} -1 \\ 4 \end{pmatrix}, \quad \ldots$$
Tenemos que la sucesión de la primera coordenada $(x_{k,1})_k$ está dada por $x_{1,k} = 1-k$.

2. Podemos considerar la sucesión $(y_i)_{i\in\mathbb{N}} \subseteq \mathbb{R}^3$ dada por
$$y_i = \left(\frac{\cos(i)}{i^4}, \frac{i^3}{i!}, \frac{1-e^{-i}}{i^4}\right).$$
Tenemos que la sucesión de la tercera coordenada $(y_{k,3})_k$ está dada por $y_{k,3} = \frac{1-e^{-k}}{k^4}$. Notar que en este ejemplo cambiamos el índice al tomar la tercera coordenada (de i a k), pero todavía estamos describiendo correctamente la misma sucesión.

3. Consideremos dos sucesiones reales $(a_k)_k, (b_k)_k \subseteq \mathbb{R}$. Podemos construir una sucesión $(z_k) \subseteq \mathbb{R}^4$ dada por
$$z_k = \left(a_k, b_k, \frac{a_k b_k}{a_k + b_k^2}, \frac{b_k(a_k^2 + b_k^2)^{2/3}}{(a_k^2 + b_k^2)^2 + b_k^2}\right).$$

$\Diamond$

Atención

El uso de índices en sucesiones vectoriales puede ser confuso. En este recuadro, resumimos la notación descrita anteriormente:

(x_k)	$\rightarrow$	Denota una sucesión, que distingumos por el uso de paréntesis.
x_k	$\rightarrow$	Denota el k-ésimo elemento de la sucesión.
$x_{k,i}$	$\rightarrow$	Denota la i-ésima coordenada del k-ésimo elemento de la sucesión.
$(x_{k,i})_k$	$\rightarrow$	Denota la sucesión real dada por las i-ésimas coordenadas de cada elemento de la sucesión (x_k).

Con la definición de sucesión vectorial construida, estudiaremos ahora la noción de límites para estas. El objetivo será definir el límite de sucesiones de tal manera que podamos **reducir el análisis al caso de una dimensión**. Para esto, lo ideal es definir el límite de una sucesión vectorial en base a las sucesiones de sus coordenadas.

Definición 1.3.3 (Convergencia de Sucesiones Vectoriales)**.** *Decimos que una sucesión $(x_k)_k$ en $\mathbb{R}^n$ **converge** a un vector $\bar{x} \in \mathbb{R}^n$, si para toda coordenada $i \in [n]$, la sucesión $(x_{i,k})_k$ de i-ésima coordenada converge (como sucesión real) a la i-ésima coordenada de $\bar{x}$. Es decir, se tiene que*

$$(x_k)_k \text{ converge a } \bar{x} \iff \forall i \in [n],\, x_{k,i} \xrightarrow{k\to\infty} \bar{x}_i. \tag{1.25}$$

*En caso de que lo anterior se cumpla, $\bar{x}$ se denomina **límite** de $(x_k)_k$ y escribimos*

$$\lim_{k\to\infty} x_k = \bar{x} \quad \text{y} \quad x_k \xrightarrow{k\to\infty} \bar{x}.$$

La noción de convergencia dada por la Definición 1.3.3 está íntimamente relacionada con la distancia de la sucesión a su límite: Una sucesión (x_k) en $\mathbb{R}^n$ converge a un punto $\bar{x} \in \mathbb{R}^n$ si, intuitivamente, la sucesión (x_k) se va acercando a $\bar{x}$ en la medida que $k \to \infty$. En el caso de una dimensión, esta idea es formalizada en la Definición 1.2.3. El siguiente teorema, muestra que podemos dar varias definiciones equivalentes en el caso vectorial.

Teorema 1.3.4 (Caracterización métrica de convergencia de sucesiones vectoriales)**.** *Sea (x_k) una sucesión en $\mathbb{R}^n$ y $\bar{x} \in \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:*

(a) $x_k \xrightarrow{k\to\infty} \bar{x}$.

(b) $d(x_k, \bar{x}) \xrightarrow{k\to\infty} 0$.

(c) $(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)\ \|x_k - \bar{x}\| \leq \varepsilon$.

Demostración. Demostraremos el ciclo de implicancias $(a) \implies (b) \implies (c) \implies (a)$.

$(a) \implies (b)$**:**

- Como $x_k \to \bar{x}$, tenemos que para toda coordenada i, $x_{k,i} \to \bar{x}_i$
- Ocupando la Proposición 1.2.9, tenemos que $\forall i \in [n], (x_{k,i} - \bar{x}_i) \to 0$.

- Ocupando la Propiedad 1 de 1.2.8, tenemos que $\forall i \in [n]$, $|x_{k,i} - \bar{x}_i| \to 0$.
- Ocupando que la función $f(t) = t^2$ es continua, y aplicando la Proposición 1.2.11, tenemos que $\forall i \in [n]$, $|x_{k,i} - \bar{x}_i|^2 \to 0^2 = 0$.
- Usando la suma de sucesiones nulas (Propiedad 3 de 1.2.8), tenemos que $\sum_{i=1}^{n} |x_{k,i} - \bar{x}_i|^2 \to 0$.
- Finalmente, ocupando que la función $f(t) = \sqrt{|t|}$ es continua, nuevamente por la Proposición 1.2.11, tenemos que

$$d(x_k, \bar{x}) = \sqrt{\sum_{i=1}^{n} |x_{k,i} - \bar{x}_i|^2} \to \sqrt{0} = 0.$$

$(b) \implies (c)$**:**

- Tenemos que $d(x_k, \bar{x}) \to 0$.
- Por definición de distancia: $\|x_k - \bar{x}\| \to 0$.
- Definiendo $r_k = \|x_k - \bar{x}\|$, tenemos que (r_k) es nula.
- Usando la Definición 1.2.6: $\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}, \forall k \geq k_0,\ |r_k| \leq \varepsilon$.
- Reemplazando tenemos que $\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}, \forall k \geq k_0,\ \|x_k - \bar{x}\| \leq \varepsilon$, que es lo que queríamos demostrar.

$(c) \implies (a)$**:**

- Sea $i \in [n]$ arbitrario. Queremos demostrar que $(x_{k,i})$ converge a $\bar{x}_i$.
- Usando (c) sabemos que $\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}, \forall k \geq k_0,\ \|x_k - \bar{x}\| \leq \varepsilon$.
- Notemos que $|x_{k,i} - \bar{x}_i| = \sqrt{|x_{k,i} - \bar{x}_i|^2} \leq \sqrt{\sum_{j=1}^{n} |x_{k,j} - \bar{x}_j|^2} = \|x_k - \bar{x}\|$.
- Entonces en particular:

$$\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}, \forall k \geq k_0,\ |x_{k,i} - \bar{x}_i| \leq \varepsilon,$$

  que es la definición de $x_{k,i} \to \bar{x}$.
- Como la coordenada $i \in [n]$ es arbitraria, concluimos que para todo $i \in [n]$, se tiene que $x_{k,i} \to \bar{x}_i$. Es decir, concluimos que $x_k \to \bar{x}$.

□

Dado que la convergencia de sucesiones vectoriales es comparable a la convergencia de sucesiones reales, es de esperar que las propiedades principales que se tenían para sucesiones reales se preserven en el caso multidimensional.

Proposición 1.3.5 (Unicidad del límite de sucesiones vectoriales)**.** *Sea $(x_k)_{k\in\mathbb{N}}$ una sucesión en $\mathbb{R}^n$. Si el límite de $(x_k)_{k\in\mathbb{N}}$ existe, entonces es único.*

Demostración. Supongamos por contradicción que $(x_k)_{k\in\mathbb{N}}$ admite dos límites distintos, u y v en $\mathbb{R}^n$. Sea $\varepsilon > 0$. Ocupando el Teorema 1.3.4, sabemos que existen $k_1, k_2 \in \mathbb{N}$ tales que

$$\forall k \geq k_1, \|x_k - u\| \leq \frac{\varepsilon}{2} \quad \text{y} \quad \forall k \geq k_2, \|x_k - v\| \leq \frac{\varepsilon}{2}.$$

Tomando $k_0 = \max\{k_1, k_2\}$, podemos escribir

$$\|u - v\| = \|u - x_{k_0} + x_{k_0} - v\| \leq \|u - x_{k_0}\| + \|x_{k_0} - v\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Como $\varepsilon > 0$ es arbitrario, tenemos que para todo $\varepsilon > 0$, se tiene que $\|u - v\| \leq \varepsilon$. Esto solo puede ocurrir si $\|u - v\| = 0$. Ocupando la propiedad de separación de la norma (ver Proposición 1.1.4), concluimos que $u = v$, lo cual es una contradicción. Concluimos entonces que, de existir, el límite de (x_k) tiene que ser único. $\square$

Proposición 1.3.6 (Álgebra de sucesiones vectoriales)**.** *Sean $(x_k)_{k\in\mathbb{N}}, (y_k)_{k\in\mathbb{N}}$ dos sucesiones en $\mathbb{R}^n$ tal que $\lim x_k = \bar{x}$ y $\lim y_k = \bar{y}$. Para todo $\alpha, \beta \in \mathbb{R}$ se tiene que*

$$\lim_k (\alpha x_k + \beta y_k) = \alpha\bar{x} + \beta\bar{y}.$$

Demostración.

- Fijemos una coordenada $i \in [n]$.
- Tenemos que $x_{k,i} \to \bar{x}_i$ y que $y_{k,i} \to \bar{y}_i$.
- Ocupando álgebra de límites de sucesiones reales (ver Proposición 1.2.11), concluimos que

$$(\alpha x_k + \beta y_k)_i = \alpha x_{k,i} + \beta y_{k,i} \xrightarrow{k\to\infty} \alpha\bar{x}_i + \beta\bar{y}_i.$$

- Como la coordenada $i \in [n]$ es arbitraria, lo anterior ocurre para todas las coordenadas. Es decir, concluimos que

$$\alpha x_k + \beta y_k \xrightarrow{k\to\infty} \alpha\bar{x} + \beta\bar{y}.$$

$\square$

Proposición 1.3.7 (Teorema del Sándwich para sucesiones vectoriales)**.** *Sean $(x_k)_{k\in\mathbb{N}}$ e $(y_k)_{k\in\mathbb{N}}$ dos sucesiones en $\mathbb{R}^n$, y sea $u \in \mathbb{R}^n$ tales que*

(i) Existe $k_0 \in \mathbb{N}$ tal que $\|x_k - u\| \leq \|y_k - u\|$ para todo $k \geq k_0$.

(ii) $\lim_k y_k = u \in \mathbb{R}^n$.

Se tiene entonces que la sucesión $(x_k)_k$ es convergente con $\lim_k x_k = u$.

Demostración. Definamos las sucesiones reales $(s_k), (t_k) \subseteq \mathbb{R}$ dadas por $s_k = \|x_k - u\|$ y $t_k = \|y_k - u\|$. Tenemos entonces que las hipótesis (i) y (ii) pueden ser escritas como

- Existe $k_0 \in \mathbb{N}$ tal que $0 \leq s_k \leq t_k$ para todo $k \geq k_0$.
- $\lim t_k = 0$.

Con esto, podemos aplicar el teorema del Sándwich para sucesiones reales (ver Proposición 1.2.13), considerando la sucesión constante $r_k = 0$, que acota por abajo a s_k y que claramente converge a 0. Concluimos así que $s_k \to 0$.
Luego, ocupando el Teorema 1.3.4, tenemos que

$$s_k \to 0 \iff \|x_k - u\| \to 0 \iff x_k \to u.$$

Esto finaliza la demostración. $\square$

1.4 Abiertos y Cerrados

La norma euclidiana también induce propiedades en los conjuntos de $\mathbb{R}^n$. Entender estas propiedades de conjuntos nos permitirá luego estudiar las funciones definidas sobre ellos. En el caso de $\mathbb{R}$, algunas propiedades de funciones, como diferenciabilidad, las estudiamos en intervalos *abiertos*: en un intervalo abierto, es posible estudiar las variaciones de cualquier punto en ambas direcciones. Otras propiedades, como los límites laterales y continuidad, las estudiamos en intervalos *cerrados*: en un intervalo cerrado, los bordes pertenecen al conjunto. El objetivo de esta sección es generalizar la noción de intervalos abiertos y cerrados en $\mathbb{R}$ a conjuntos (abiertos y cerrados) que preserven estas características.

1.4.1 Bolas en $\mathbb{R}^n$

Una primera idea es pensar que los intervalos son conjuntos definidos en torno a su punto medio. Sean $a, b \in \mathbb{R}$ con $a < b$, y consideremos el intervalo abierto (a, b). Tomando $m = \frac{a+b}{2}$ y $r = \frac{b-a}{2}$, podemos pensar el intervalo I como el conjunto de puntos a distancia de m menor que r, es decir,

$$(a, b) = \{x \in \mathbb{R} \;\; |x - m| < r\}. \tag{1.26}$$

la Figura 1.9 ilustra esta idea. Similarmente, el intervalo cerrado $[a, b]$ se puede definir de la misma manera, pero con desigualdad simple ($\leq$) en vez de estricta.

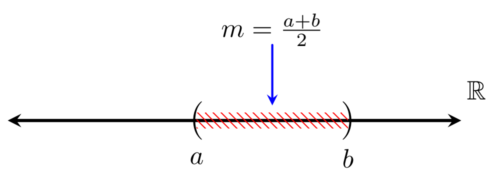


Figura 1.9: Intervalo abierto (a, b) y punto medio m.

Esta observación nos permite hacer una primera generalización: al pasar a dimensión superior, basta reemplazar el valor absoluto por la norma en la ecuación 1.26. Esto nos lleva a la siguiente definición:

Definición 1.4.1 (Bolas en $\mathbb{R}^n$)**.** *Sea $n \in \mathbb{N}$. Para un punto $x \in \mathbb{R}^n$ y un real $r > 0$, definimos*

*(a) La bola **abierta** de centro x y radio r como el conjunto*

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

*(b) La bola **cerrada** de centro x y radio r como el conjunto*

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

En el caso particular de considerar $x = 0$ y $r = 1$, llamaremos a la bola cerrada $\overline{B}(0, 1)$ la bola unitaria, y la denotamos por $\mathbb{B}$.

La Figura 1.10 ilustra la bola unitaria en dimensiones $n = 1, 2$ y 3. La gracia de la bola unitaria $\mathbb{B}$ es que, independiente del centro x o el radio r, una bola cerrada $\overline{B}(x, r)$ tiene la misma forma esférica que $\mathbb{B}$.

Las bolas abiertas y cerradas son una buena generalización de los intervalos. Sin embargo, en $\mathbb{R}^n$, los conjuntos pueden llegar a ser mucho más complejos. Pueden estar definidos por varias igualdades y/o desigualdades, e incluso pueden no tener descripción analítica.

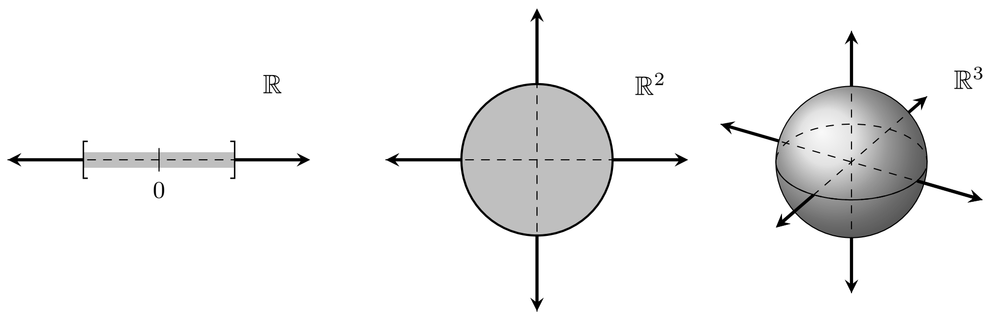


Figura 1.10: Bola unitaria $\mathbb{B}$ en $\mathbb{R}$, $\mathbb{R}^2$ y $\mathbb{R}^3$.

Ejemplo 1.4.2 (Conjuntos en $\mathbb{R}^n$)**.** Algunos ejemplos de conjuntos en $\mathbb{R}^2$ o $\mathbb{R}^3$ son:

1. $A = \{(x, y, z) \in \mathbb{R}^3 \ : \ z = 1, y > \frac{1}{5}x^2, x \geq 0\}$. Este conjunto corresponde a todos los puntos de $\mathbb{R}^3$ en en el plano $[z = 1]$, cuya segunda coordenada es estrictamente mayor que el cuadrado de la primera, y cuya primera coordenada es mayor o igual a cero. Notar que este conjunto está definido únicamente por igualdades y desigualdades. La Figura 1.11 ilustra el conjunto A.

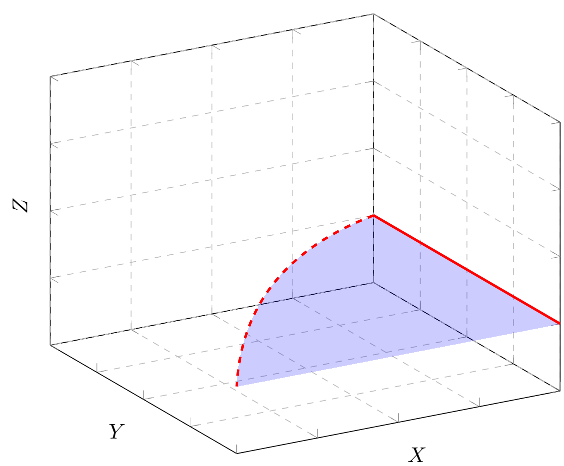


Figura 1.11: Dibujo del conjunto A (la parte comprendida en el cubo $[0, 5]^3$). Las lineas rojas representan el borde del conjunto. La linea continua $\{x = 0, z = 1\}$ pertenece al conjunto, mientras que la línea punteada $\{y = \frac{1}{5}x^2, z = 1\}$, no.

2. $B = \{(x, y) \in \mathbb{R}^2 \ : \ y > x^2 - 2, y < |x|\} \cup \{(0, 1)\}$. Este conjunto corresponde a todos los puntos en $(x, y) \in \mathbb{R}^2$, que quedan por sobre el gráfico de la parábola $y = x^2 - 1$ y estrictamente por debajo del gráfico de $y = |x|$, incluyendo a parte el punto $(0, 1)$. Notar que este conjunto está definido por igualdades y desigualdades, y por álgebra de conjuntos (unión, intersección, complemento). La Figura 1.12 ilustra el conjunto A.

3. Los conjuntos podrían no estar descritos de forma analítica. El conjunto C de la Figura 1.13 es un caso.

$\Diamond$

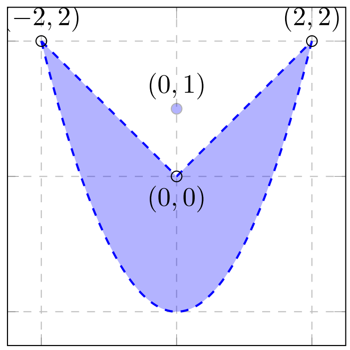


Figura 1.12: Dibujo del conjunto B. La línea punteada no pertenece al conjunto, al igual que los vértices (-2,2), (0,0) y (2,2). El punto (0,1) es incluido en el conjunto vía álgebra.

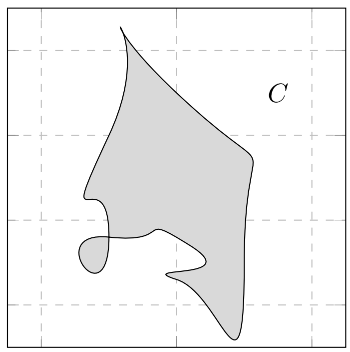


Figura 1.13: Ejemplo de conjunto no-paramétrico.

1.4.2 Puntos interiores y conjuntos abiertos

Un elemento fundamental en el desarrollo del cálculo, son las variaciones en torno a un punto del dominio de una función. Es decir, para un punto $x \in A$, queremos estudiar como se comporta una función $f : A \to \mathbb{R}$ "cerca de x".
En un intervalo abierto, esta idea se traducía en el hecho que para cualquier punto puedo "moverme un poco" en ambas direcciones sin salirme del conjunto, como ilustra la Figura 1.14

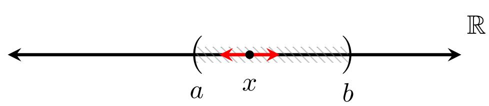


Figura 1.14: Variaciones en torno a un punto para intervalos abiertos.

Esto nos permitía, para una función $f : (a, b) \to \mathbb{R}$, calcular ambos límites laterales y eventualmente su derivada.

En $\mathbb{R}^n$, los conjuntos pueden ser mucho más extraños. Aún así, estamos interesados en los puntos donde "puedo moverme un poco **en todas las direcciones**" sin salirme del conjunto. En un conjunto A, moverse en todas las direcciones desde un punto de referencia x, equivale a que A contenga una bola centrada en x como subconjunto (ver Figura 1.15). Esto motiva la definición de *punto interior*.

Figura 1.15: Ilustración de que todas las direcciones significa tomar una bola en torno al punto. El radio de la bola puede depender del punto.

Definición 1.4.3 (Puntos interiores e interior de un conjunto)**.** *Sea A un subconjunto de $\mathbb{R}^n$. Decimos que $x \in A$ es **punto interior** de A, si existe $r > 0$ tal que*

$$B(x, r) \subseteq A.$$

*El conjunto de todos los puntos interiores de A se llama **interior de A** y se denota por $\operatorname{int}(A)$ o por $\mathring{A}$.*

Claramente, todo punto interior de un conjunto A es un punto de A, es decir, $\operatorname{int}(A) \subseteq A$. Cuando la otra inclusión se tiene, diremos que el conjunto es *abierto*.

Definición 1.4.4 (Conjunto abierto)**.** *Un conjunto $A \subseteq \mathbb{R}^n$ se dice **abierto** si $A = \operatorname{int}(A)$.*

Para entender mejor los conjuntos abiertos, es útil conocer qué conjuntos abiertos son los "básicos" y cómo los conjuntos abiertos interactúan bajo unión e intersección (las operaciones básicas asociadas a los conjuntos).

Proposición 1.4.5 (Álgebra de conjuntos abiertos)**.** *Se tiene las siguientes propiedades:*

1. *$\emptyset$ y $\mathbb{R}^n$ son conjuntos abiertos.*
2. *Para todo punto $x \in \mathbb{R}^n$, y todo radio $r > 0$, la bola abierta $B(x, r)$ es un conjunto abierto.*
3. *Para una colección finita $\{A_i\}_{i=1}^m$ de conjuntos abiertos de $\mathbb{R}^n$, se tiene que la unión $\bigcup_{i=1}^m A_i$ es un conjunto abierto.*
4. *Para una colección finita $\{A_i\}_{i=1}^m$ de conjuntos abiertos de $\mathbb{R}^n$, se tiene que la intersección $\bigcap_{i=1}^m A_i$ es un conjunto abierto.*

Demostración. Para demostrar que un conjunto es abierto, queremos ver que todo punto del conjunto es un punto interior.

1. Demostremos cada uno por separado
   - El conjunto $\emptyset$ es abierto por argumento de vacuidad (ver nota más adelante). En efecto, no hay puntos en $\emptyset$ sobre los cuales verificar si son interiores.
   - Queremos demostrar que $\mathbb{R}^n$ es un conjunto abierto.
     - Sea $x \in \mathbb{R}$ arbitrario.

- Consideremos la bola $B(x,1)$ y notemos que $B(x,1) \subseteq \mathbb{R}^n$ por definición (ver Definición 1.4.1)
- Demostramos que existe $r = 1 > 0$ tal que $B(x,r) \subseteq \mathbb{R}^n$, por lo que x es punto interior.
- Como x es arbitrario, concluimos que todo punto de $\mathbb{R}^n$ es punto interior, por lo que $\mathbb{R}^n$ es un conjunto abierto.

2. Tomemos la bola abierta $B(x,r)$. Queremos ver que es un conjunto abierto. La Figura 1.16 muestra un esquema en $\mathbb{R}^2$ de los elementos que utilizan en la demostración.

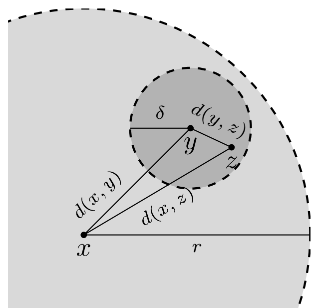


Figura 1.16: Representación en $\mathbb{R}^2$ de los elementos de la demostración de $B(x,r)$ es abierta.

- Sea $y \in B(x,r)$ un punto cualquiera (arbitrario).
- Como $y \in B(x,r)$, entonces $d(x,y) < r$.
- Tomemos $\delta > 0$ suficientemente pequeño tal que $d(x,y) + \delta < r$ (1). Por ejemplo, se podría tomar $\delta = \frac{r-d(x,y)}{2} > 0$.
- Queremos ver que $B(y,\delta) \subseteq B(x,r)$. Para eso debemos tomar un punto cualquiera en $B(y,\delta)$ y demostrar que está en $B(x,r)$
  - Sea $z \in B(y,\delta)$ arbitrario.
  - Por definición: $z \in B(x,r) \iff d(x,z) < r$. En efecto,

$$\begin{aligned} d(x,z) &= \|x - z\| \\ &= \|x - y + y - z\| \\ &\leq \|x - y\| + \|y - z\| \\ &= d(x,y) + d(y,z) \qquad (1) \\ &< d(x,y) + \delta \\ &< r. \end{aligned}$$

    Concluimos que $x \in B(x,r)$.
  - Como el punto z es arbitrario, tenemos que $B(y,\delta) \subseteq B(x,r)$.
- Demostramos que existe $\delta > 0$ tal que $B(y,\delta) \subseteq B(x,r)$, es decir, y es punto interior de $B(x,r)$.
- Como el punto y es arbitrario, concluimos que todo punto en $B(x,r)$ es punto interior.
- Tenemos que $B(x,r)$ es un conjunto abierto.

3. Tenemos una colección finita $\{A_i\}_{i=1}^m$ de conjuntos. Queremos demostrar que $\bigcup_{i=1}^m A_i$ es un conjunto abierto.

- Sea $x \in \bigcup_{i=1}^m A_i$ arbitrario.

- Como x está en la unión, entonces necesariamente está en alguno de los conjuntos, es decir, $x \in A_{\bar{\imath}}$ para algún $\bar{\imath} \in [m]$.
- Como $A_{\bar{\imath}}$ es un conjunto abierto, sabemos que existe $r > 0$ tal que $B(x, r) \subseteq A_{\bar{\imath}}$.
- Por definición de la unión:

$$B(x, r) \subseteq A_{\bar{\imath}} \subseteq \bigcup_{i=1}^{m} A_i.$$

- Entonces hemos demostrado que existe $r > 0$ tal que $B(x, r) \subseteq \bigcup_{i=1}^{m} A_i$, por lo que x es un punto interior.
- Como x es arbitrario, concluimos que todos los puntos de $\bigcup_{i=1}^{m} A_i$ son interiores, es decir, $\bigcup_{i=1}^{m} A_i$ es un conjunto abierto.

4. Tenemos una colección finita $\{A_i\}_{i=1}^{m}$ de conjuntos. Queremos demostrar que $\bigcap_{i=1}^{m} A_i$ es un conjunto abierto.

   - Sea $x \in \bigcap_{i=1}^{m} A_i$ arbitrario.
   - Como x está en la intersección, entonces necesariamente está en todos de los conjuntos, es decir, $x \in A_i$ para todo $i \in [m]$.
   - Para cada $i \in [m]$, A_i es un conjunto abierto, por lo que existe $r_i > 0$ tal que $B(x, r_i) \subseteq A_i$.
   - Tomemos $\bar{r} = \text{mín}\{r_1, \ldots, r_m\}$. Como es el mínimo entonces $\bar{r} \leq r_i$ para todo $i \in [m]$ y, en consecuencia, $B(x, \bar{r}) \subseteq B(x, r_i)$ para todo $i \in [m]$.
   - Concluimos que $B(x, \bar{r}) \subseteq \bigcap_{i=1}^{m} A_i$, por lo que x es punto interior.
   - Como x es arbitrario, demostramos entonces que todos los puntos de $\bigcap_{i=1}^{m} A_i$ son interiores, es decir, que $\bigcap_{i=1}^{m} A_i$ es un conjunto abierto.

$\square$

Nota: Argumento de vacuidad

El argumento de vacuidad en lógica establece que la proposición cuantificada

$$\forall x \in \emptyset, P(x)$$

es siempre verdadera, independiente de cual sea la función proposicional P. Una forma de ver esto es que, al no haber elementos en $\emptyset$, la proposición $P(x)$ nunca será evaluada y por lo tanto, nunca será falsa. Es decir, la negación

$$\exists x \in \emptyset, \neg P(x)$$

es siempre falsa, simplemente pues no existen elementos en $\emptyset$.

1.4.3 Puntos adherentes y conjuntos cerrados

Otro concepto central del cálculo, al momento de estudiar límites en un conjunto $A \subseteq \mathbb{R}^n$ (tanto de sucesiones, en la sección 1.3, como de funciones, en la sección 1.5), necesitamos conocer cuales son los puntos a los que nos podemos aproximar desde A. Intuitivamente, un punto en $\mathbb{R}^n$ que esté lejos del conjunto A no podrá ser aproximado. Por otro lado, los puntos pertenecientes al conjunto A así como los puntos que están "a distancia cero" del conjunto A, si pueden ser aproximados. En la Figura 1.17, se ilustra la idea.
Para formalizar esta idea, introducimos la definición de puntos adherentes a un conjunto, precisamente como aquellos puntos que que están "infinitamente cerca" o a "distancia cero".

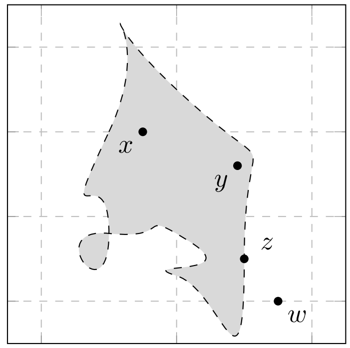


Figura 1.17: Los puntos x e y están en el conjunto A, por lo que si pueden ser aproximados. Asimismo, el punto z no está en el conjunto A, pero está "a distancia cero", por lo que si puede ser aproximado desde A. Finalmente, el punto w no puede ser aproximado desde A, pues "está lejos".

Definición 1.4.6 (Puntos adherentes y adherencia de un conjunto)**.** *Sea $A \subseteq \mathbb{R}^n$ un conjunto cualquiera. Diremos que un punto $\bar{x} \in \mathbb{R}^n$ es* ***punto adherente*** *de A si para todo $r > 0$, la bola $B(\bar{x}, r)$ intersecta al conjunto A, es decir, si se verifica:*

$$\forall r > 0,\ B(x, r) \cap A \neq \emptyset \tag{1.27}$$

El conjunto de todos los puntos adherentes de A se llama ***adherencia de A****, y se denota por $\mathrm{adh}(A)$ o por $\overline{A}$.*

En la Figura 1.17, tenemos que x, y y z son puntos adherentes, mientras que el punto w no lo es. La siguiente proposición, nos entrega algunas caracterizaciones de la adherencia de un conjunto.

Proposición 1.4.7 (Caracterizaciones de adherencia)**.** *Sea $A \subseteq \mathbb{R}^n$ y sea $x \in \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:*

(a) $x \in \mathrm{adh}(A)$.

(b) *Existe una sucesión $(x_n) \subseteq A$ tal que $x_n \to x$.*

(c) *Se tiene que $\inf\{\|z - x\| \ :\ z \in A\} = 0$.*

Demostración. Demostraremos el ciclo de implicancias $(a) \implies (b) \implies (c) \implies (a)$.

$(a) \implies (b)$**:** Para esta demostración primero vamos a construir una <u>sucesión en A</u> que "debería" converger a x para luego verificar que efectivamente converge a x.

- Como $x \in \mathrm{adh}(A)$, para todo $r > 0$: $B(x, r) \cap A \neq \emptyset$.
- En particular, para cada $k \in \mathbb{N}$, tomando $r = \frac{1}{k} > 0$, tenemos que $B\left(x, \frac{1}{k}\right) \cap A \neq \emptyset$.
- Como la intersección es no vacía, tomemos para cada $k \in \mathbb{N}$, $x_k \in B\left(x, \frac{1}{k}\right) \cap A$.
- Hemos construido una sucesión $(x_k)_{k\in\mathbb{N}} \subseteq A$. Veamos que $x_k \to x$:
  - Por definición de distancia $0 \leq d(x, x_k)$.
  - Como $x_k \in B\left(x, \frac{1}{k}\right)$, entonces $d(x, x_k) < \frac{1}{k}$.
  - Aplicando el Teorema del Sándwich (ver Proposición 1.2.13) tenemos que que $d(x, x_k) \to 0$.
  - Dada la caracterización métrica de convergencia (Teorema 1.3.4), concluimos que $x_k \to x$.

$(b) \implies (c)$**:** Llamemos $\alpha = \text{ínf}\{\|z-x\| \ : \ z \in A\}$.

- Primero, sabemos que $\alpha \geq 0$ ya que es el ínfimo de valores positivos.
- Queremos demostrar que $\alpha = 0$. Razonando por contradicción, supongamos $\alpha > 0$.
- Tenemos que $x_n \to x$, es decir, $(\forall \varepsilon > 0)(\exists k_0 \in \mathbb{N})(\forall k \geq k_0)\ \|x_k - x\| \leq \varepsilon$.
- Tomando $\varepsilon = {}^{\alpha}/_{2} > 0$, definamos $z = x_{k_0}$ con $k_0 \in \mathbb{N}$ del punto anterior.
- Concluimos que $\exists z \in A : \ \|z - x\| \leq \varepsilon < \alpha$. Lo que es una contradicción, pues α es el ínfimo de dichas normas.
- Por lo tanto, necesariamente $\alpha = \text{ínf}\{\|z-x\| \ : \ z \in A\} = 0$.

$(c) \implies (a)$**:** Recordemos que el ínfimo es la mayor de las cotas inferiores de un conjunto en $\mathbb{R}$.

- Sea $r > 0$ cualquiera, entonces $\text{ínf}\{\|z-x\| \ : \ z \in A\} = 0 < r$.
- Existe $z \in A$ tal que $\|z - x\| \leq r$ (si sno el ínfimo sería mayor o igual a r).
- Tenemos que $z \in A$ y $z \in B(x,r)$, por lo tanto $z \in A \cap B(x,r)$.
- Como r es arbitrario, tenemos que $\forall r > 0,\ \exists z \in A \cap B(x,r)$
- De forma equivalente: $\forall r > 0, A \cap B(x,r) \neq \emptyset$.
- Es decir, $x \in \text{adh}(A)$

□

La caracterización (c) en la Proposición 1.4.7 motiva introducir la función distancia a un conjunto A, como el ínfimo de las distancias a los puntos de A. Es decir, definimos la función distancia a un conjunto A no-vacio como

$$\begin{aligned} d_A : \mathbb{R}^n &\to \mathbb{R}_+ \\ x &\mapsto d_A(x) := \text{ínf}\{\|z-x\| \ : \ z \in A\}. \end{aligned} \tag{1.28}$$

Esta definición nos permite computar que tan lejos está un punto de A, y la propiedad (c) se puede escribir como

$$x \in \text{adh}(A) \iff d_A(x) = 0. \tag{1.29}$$

Atención

Esta definición de función distancia $d_A(\cdot)$ necesita que A sea no-vacía. De lo contrario, el ínfimo de la ecuación (1.28) estaría mal definido (sería $+\infty$). No obstante, admitiendo la convención $d_\emptyset \equiv +\infty$, entonces la caracterización (1.29) es siempre válida.

Como hemos dicho hasta ahora, para todo conjunto A, cada vector $x \in A$ es punto adherente de A. Esto nos permite concluir que $A \subseteq \text{adh}(A)$. Cuando la otra inclusión se tiene, diremos que el conjunto A es *cerrado*.

Definición 1.4.8 (Conjuntos cerrados)**.** *Sea $C \subseteq \mathbb{R}^n$. Diremos que C es un conjunto* ***cerrado*** *si $C = \text{adh}(C)$.*

Nota: Cerradura

El concepto de conjunto cerrado es un concepto fundamental de topología general. Por lo mismo, muchos libros llaman a la adherencia de un conjunto como su *cerradura* o *clausura*. Consistentemente, también se ocupa la notación $\mathrm{cl}(A)$ para referirse a la adherencia de A.

Proposición 1.4.9 (Caracterización secuencial de conjuntos cerrados)**.** *Sea $C \subseteq \mathbb{R}^n$. Se tiene que*

$$C \text{ es cerrado} \iff \forall (x_k)_k \subseteq C \text{ convergente}, \lim x_k \in C. \tag{1.30}$$

Demostración. Mostraremos la doble implicancia.

($\implies$)

- Supongamos que C es cerrado (hipótesis).
- Sea $(x_k) \subseteq C$ una sucesión convergente cualquiera y llamemos $\bar{x} = \lim_k x_k$.
- Por la Proposición 1.4.7, tenemos que $\bar{x} \in \mathrm{adh}(C)$.
- Como $\mathrm{adh}(A) = A$ (por hipótesis de cerradura), concluimos que $\bar{x} \in A$.
- Como la sucesión es arbitraria, demostramos que

$$\forall (x_k)_k \subseteq C \text{ convergente}, \lim x_k \in C.$$

($\impliedby$)

- Supongamos que para toda sucesión (x_k) en C convergente, el límite $\lim_k x_k \in C$ (hipótesis).
- Sea $\bar{x} \in \mathrm{adh}(C)$ cualquiera.
- Por la Proposición 1.4.7, sabemos que existe $(x_n) \subseteq C$ tal que $x_n \to \bar{x}$.
- Dada la hipótesis, $\bar{x} \in C$.
- Como $\bar{x} \in \mathrm{adh}(C)$ es arbitrario, tenemos que $\mathrm{adh}(C) \subseteq C$.
- Concluimos que $\mathrm{adh}(C) = C$, es decir, C es cerrado.

□

Los conjuntos abiertos y los conjuntos cerrados están íntimamente relacionados, pues los abiertos son los complementos de los cerrados (y viceversa).

Teorema 1.4.10 (Relación entre abiertos y cerrados)**.** *Sea $A \subseteq \mathbb{R}^n$. Se tiene que*

$$A \text{ es abierto} \iff A^c \text{ es cerrado.}$$

Demostración. Mostraremos la doble implicancia:

($\Longrightarrow$)

- Asumamos que A es abierto (hipótesis).
- Para ver que A^c es cerrado, usemos la caracterización de la Proposición 1.4.9.
- Sea $(x_k) \subseteq A^c$ una sucesión convergente cualquiera, con límite $\bar{x} = \lim x_k$.
- Razonando por contradicción, supongamos que $\bar{x} \notin A^c$, es decir, $\bar{x} \in A$.
    - Como A es abierto, existe $r > 0$ tal que $B(x, r) \subseteq A$.
    - Como $x_k \to \bar{x}$, tomando $\varepsilon = \frac{r}{2}$ en el Teorema 1.3.4 sabemos que existe k_0 tal que $\|x_{k_0} - \bar{x}\| \leq \varepsilon < r$
    - Por lo tanto $x_{k_0} \in B(x, r) \subseteq A$ y $x_{k_0} \in A^c$ por ser parte de la sucesión, lo que es una contradicción.
- Concluimos que $\bar{x} \in A^c$ y, en consecuencia, A^c es cerrado.

($\Longleftarrow$)

- Asumamos que A^c es cerrado, es decir, $\mathrm{adh}(A^c) = A^c$ (hipótesis).
- Sea $x \in A$ cualquiera, veamos que es punto interior.
- Llamemos $\alpha = d_{A^c}(x) = \inf\{\|z - x\| \ : \ z \notin A\}$
- Notemos que $\alpha > 0$ (si es cero, por (1.29) significa que $x \in \mathrm{adh}(A^c) = A^c$).
- Por definición del ínfimo la bola $B(x, \alpha) = \{z \in \mathbb{R}^n : \|z - x\| < \alpha\} \subseteq A$.
- Concluimos que x es punto interior de A, y al ser arbitrario demostramos que A es abierto.

$\square$

Proposición 1.4.11 (Álgebra de conjuntos cerrados)**.** *Se tiene las siguientes propiedades:*

1. *$\emptyset$ y $\mathbb{R}^n$ son conjuntos cerrados.*
2. *Para todo punto $x \in \mathbb{R}^n$, y todo radio $r > 0$, la bola cerrada $\overline{B}(x, r)$ es un conjunto cerrado.*
3. *Para una colección finita $\{C_i\}_{i=1}^m$ de conjuntos cerrados de $\mathbb{R}^n$, se tiene que la unión $\bigcup_{i=1}^m C_i$ es un conjunto cerrado.*
4. *Para una colección finita $\{C_i\}_{i=1}^m$ de conjuntos cerrados de $\mathbb{R}^n$, se tiene que la intersección $\bigcap_{i=1}^m C_i$ es un conjunto cerrado.*

Demostración. Para esta demostración usaremos constantemente el teorema anterior.

1.
    - Como $\mathbb{R}^n$ es abierto, entonces $\emptyset = (\mathbb{R}^n)^c$ es cerrado.
    - Como $\emptyset$ es abierto, entonces $\mathbb{R}^n = \emptyset^c$ es cerrado.
2. Tomemos la bola cerrada $\overline{B}(x, r)$. Demostremos que es un conjunto cerrado usando la Propiedad 1.4.9
    - Sea $(y_k) \subseteq \overline{B}(x, r)$ una sucesión convergente cualquiera. Llamemos $y = \lim_k y_k$. Queremos ver que $y \in \overline{B}(x, r)$
    - Por desigualdad triangular para distancias (ver Proposición 1.1.6, tenemos que para cada $k \in \mathbb{N}$:
    $$d(x, y) \leq d(x, y_k) + d(y_k, y) \leq r + d(y_k, y) \quad (1)$$

- Como $y_k \to y$, tenemos que $d(y_k, y) \to 0$ (ver Teorema 1.3.4)
- Tomando límite en (1), tenemos que $d(x, y) \leq r$.
- Concluimos que $y \in \overline{B}(x, r)$.
- Como (y_k) es una sucesión arbitraria, tenemos que $\overline{B}(x, r)$ es un conjunto cerrado.

3. Sea una colección finita $\{C_i\}_{i=1}^m$ de conjuntos cerrados de $\mathbb{R}^n$.
   - Los conjuntos $A_i = C_i^c$, con $i \in [m]$, son abiertos.
   - El conjunto $\left(\bigcup_{i=1}^m C_i\right)^c = \bigcap_{i=1}^m A_i$ es abierto.
   - Por lo tanto, $\bigcup_{i=1}^m C_i$, es cerrado.
4. Sea una colección finita $\{C_i\}_{i=1}^m$ de conjuntos cerrados de $\mathbb{R}^n$.
   - Los conjuntos $A_i = C_i^c$, con $i \in [m]$, son abiertos.
   - El conjunto $\left(\bigcap_{i=1}^m C_i\right)^c = \bigcup_{i=1}^m A_i$ es abierto.
   - Por lo tanto, $\bigcap_{i=1}^m C_i$, es cerrado.

□

Atención

El conjunto vacío $\emptyset$ y el espacio completo $\mathbb{R}^n$ son los únicos dos conjuntos que son abiertos y cerrados al mismo tiempo.

1.4.4 Frontera de un conjunto

De la discusión de adherencia y de interior, sabemos que para todo conjunto $A \subseteq \mathbb{R}^n$, siempre se cumple que

$$\operatorname{int}(A) \subseteq A \subseteq \operatorname{adh}(A). \tag{1.31}$$

En general, ambas inclusiones son estrictas, siendo la mayoría de los conjuntos ni abiertos ni cerrados. Para identificar si un conjunto es abierto, cerrado, o ninguna de las anteriores, la clave está en estudiar los puntos que están en $\operatorname{adh}(A) \setminus \operatorname{int}(A)$:

1. si todos estos puntos están en A, el conjunto A es cerrado;
2. si todos estos puntos están en A^c, el conjunto A es abierto;
3. si los puntos de $\operatorname{adh}(A) \setminus \operatorname{int}(A)$ están repartidos entre A y A^c, entonces A no es ni abierto ni cerrado.

Este último caso se ilustra en la Figura 1.18, y motiva la definición de frontera.

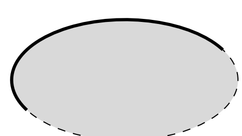
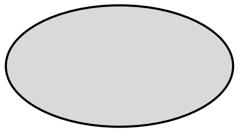


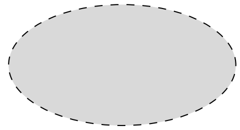
Figura 1.18: Ilustración de un conjunto que no es ni abierto ni cerrado.

Definición 1.4.12 (Frontera). *Sea A un conjunto de $\mathbb{R}^n$. Definimos la frontera de A, denotada por* $\operatorname{Fr}(A)$, *como*

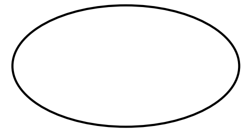
$$\operatorname{Fr}(A) = \operatorname{adh}(A) \setminus \operatorname{int}(A). \tag{1.32}$$



(a) Adherencia



(b) Interior



(c) Frontera

Figura 1.19: Ilustración de Adherencia, interior y frontera de conjuntos. Notar que cualquier conjunto entre (a) y (b) tendrá la misma adherencia, el mismo interior, y la misma frontera.

La Figura 1.19 ilustra como se relacionan los conceptos de interior, adherencia y frontera, tomando como base, el conjunto de la Figura 1.18.
La frontera de un conjunto puede interpretarse como el límite geométrico que separa el conjunto de su complemento.

Proposición 1.4.13 (Propiedades de la Frontera)**.** *Sea A un subconjunto de $\mathbb{R}^n$ y $x \in \mathbb{R}^n$. Se tiene que*

1. *El vector x pertenece a* $\mathrm{Fr}(A)$ *si y solo si se cumple que*

$$\forall r > 0, \quad B(x,r) \cap A \neq \emptyset \ \ y \ \ B(x,r) \cap A^c \neq \emptyset. \tag{1.33}$$

2. $\mathrm{Fr}(A) = \mathrm{adh}(A) \cap \mathrm{adh}(A^c)$.

3. $\mathrm{Fr}(A) = \mathrm{Fr}(A^c)$.

Demostración.

1. - De la Proposición 1.4.7, sabemos que $x \in \mathrm{adh}(A) \iff \forall r > 0,\ B(x,r) \cap A \neq \emptyset$.
   - De la Definición 1.4.3, sabemos que $x \notin \mathrm{int}(A) \iff \forall r > 0,\ B(x,r) \not\subseteq A$.
   - Esto equivale a que $x \notin \mathrm{int}(A) \iff \forall r > 0,\ B(x,r) \cap A^c \neq \emptyset$.
   - Con esto en mente, podemos escribir

$$\begin{aligned} x \in \mathrm{Fr}(A) &\iff x \in \mathrm{adh}(A) \setminus \mathrm{int}(A) \\ &\iff x \in \mathrm{adh}(A) \quad \text{y} \quad x \notin \mathrm{int}(A) \\ &\iff [\forall r > 0,\ B(x,r) \cap A \neq \emptyset] \quad \text{y} \quad [\forall r > 0,\ B(x,r) \cap A^c \neq \emptyset] \\ &\iff \forall r > 0, \quad B(x,r) \cap A \neq \emptyset \ \text{y} \ B(x,r) \cap A^c \neq \emptyset. \end{aligned}$$

2. Usando la Proposición 1.4.7 y la ecuación obtenida en el punto aterior, tenemos que:

$$\begin{aligned} x \in \mathrm{Fr}(A) &\iff [\forall r > 0,\ B(x,r) \cap A \neq \emptyset] \quad \text{y} \quad [\forall r > 0,\ B(x,r) \cap A^c \neq \emptyset] \\ &\iff x \in \mathrm{adh}(A) \quad \text{y} \quad x \in \mathrm{adh}(A^c) \\ &\iff x \in \mathrm{adh}(A) \cap \mathrm{adh}(A^c). \end{aligned}$$

3. Como $A = (A^c)^c$, tenemos que:

$$\mathrm{Fr}(A) = \mathrm{adh}(A) \cap \mathrm{adh}(A^c) = \mathrm{adh}((A^c)^c) \cap \mathrm{adh}(A^c) = \mathrm{Fr}(A^c)$$

$\square$

1.5 Límites y Continuidad de Funciones en $\mathbb{R}^n$

Toda la construcción que hemos hecho hasta ahora (norma, producto interno, sucesiones, conjuntos abiertos y cerrados) nos permitirá, de ahora en adelante, entrar en el principal objeto de estudio de este curso: las funciones de varias variables. Recordemos que en general, una función $f : A \to B$ (con A y B conjuntos cualquiera) es una asignación que a cada elemento $a \in A$, asigna un único elemento $f(a) \in B$.

En los cursos de cálculo de una variable, estudiamos en profundidad las funciones $f : (a, b) \to \mathbb{R}$, que a cada elemento real $x \in (a, b)$, entregaban un único número real $f(x) \in \mathbb{R}$. Polinomios de la forma $p(x) = a_m x^m + \cdots + a_1 x + a_0$, funciones trigonométricas como $f(x) = \sin(x)$ o $f(x) = \cosh(x)$, funciones exponencial $f(x) = e^x$ y logaritmo $f(x) = \ln(x)$ son algunos de los ejemplos fundamentales de funciones de una variable. Para estudiar propiedades de f, identificábamos el argumento de f como la variable, que usábamos para calcular límites, derivadas e integrales.

En este curso, estudiaremos funciones que reciben varias variables, y entregan varios valores reales. Dicho de otro modo, estudiaremos funciones que reciben y entregan vectores.

Definición 1.5.1 (Función vectorial)**.** *Sean $n, m \in \mathbb{N}$ y sea $\Omega \subseteq \mathbb{R}^n$ un conjunto no-vacío. Una función vectorial es una asignación $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$, que a cada vector $x = (x_1, \ldots, x_n) \in \Omega$, asigna un único vector $f(x) = (f_1(x), \ldots, f_m(x)) \in \mathbb{R}^m$.*
En tal caso, diremos que

- *Ω es el dominio de f.*
- *f es una función de n variables.*
- *f es una función de m coordenadas.*

Para cada $i \in [m]$, identificamos $f_i : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ como la función i-ésima coordenada: es decir, que a cada vector $x \in \Omega$ asigna $f_i(x) \in \mathbb{R}$ (la i-ésima coordenada del vector $f(x)$).

Ejemplo 1.5.2. Consideremos $\Omega = \{x \in \mathbb{R}^2\ x_1 > 0\}$, y definamos la función:

$$\begin{aligned} f : \Omega &\to \mathbb{R}^3 \\ x = (x_1, x_2) &\mapsto \left(\ln(x_1),\ \sqrt[3]{x_1 x_2},\ \tfrac{x_2^2}{5x_1}\right). \end{aligned}$$

Esta función recibe un vector $x \in \Omega$ y entrega el vector en $\mathbb{R}^3$ dado por $f(x) = \left(\ln(x_1), \sqrt[3]{x_1 x_2}, \tfrac{x_2^2}{5x_1}\right)$. Notar que para $x \notin \Omega$, la función f no está bien definida, pues la primera coordenada se indefine si $x_1 \leq 0$ y la tercera se indefine si $x_1 = 0$.

Por otro lado, la función f tiene tres funciones coordenadas, ilustradas en la Figura 1.20:

1. $f_1 : \Omega \to \mathbb{R}$, dada por $f_1(x) = \ln(x_1)$.
2. $f_2 : \Omega \to \mathbb{R}$, dada por $f_2(x) = \sqrt[3]{x_1 x_2}$.
3. $f_3 : \Omega \to \mathbb{R}$, dada por $f_3(x) = \frac{x_2^2}{5x_1}$.

Las expresiones que forman cada una de las funciones coordenadas dependen de todas las variables de entrada, que en este caso son x_1 y x_2. Notar que, aunque la expresión de f_1 solo esté escrita en función de x_1, todavía recibe 2 variables, x_1 y x_2. Lo que pasa es que f_1 es constante respecto a variaciones de x_2. $\Diamond$

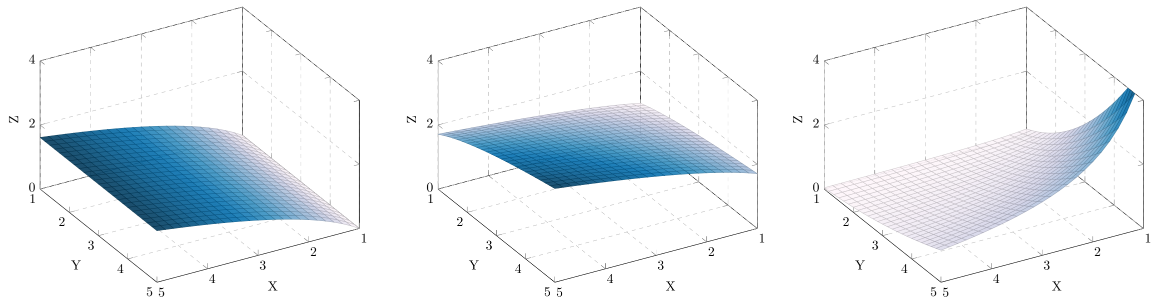


Figura 1.20: Gráfico de las tres funciones coordendas de $f(x_1, x_2) = \left(\ln(x_1), \sqrt[3]{x_1 x_2}, \frac{x_2^2}{5x_1}\right)$, ordenadas de izquierda a derecha.

Ejemplo 1.5.3. Una función puede estar definida por ramas, considerando particiones del dominio. Por ejemplo, podemos definir la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} x^2 + y^2 & \text{si } x^2 + y^2 \leq 1 \\ 1 + \frac{1}{3}\cos\left(\frac{\pi}{2}\sqrt{x^2+y^2}\right) & \text{si } x^2 + y^2 > 1. \end{cases}$$

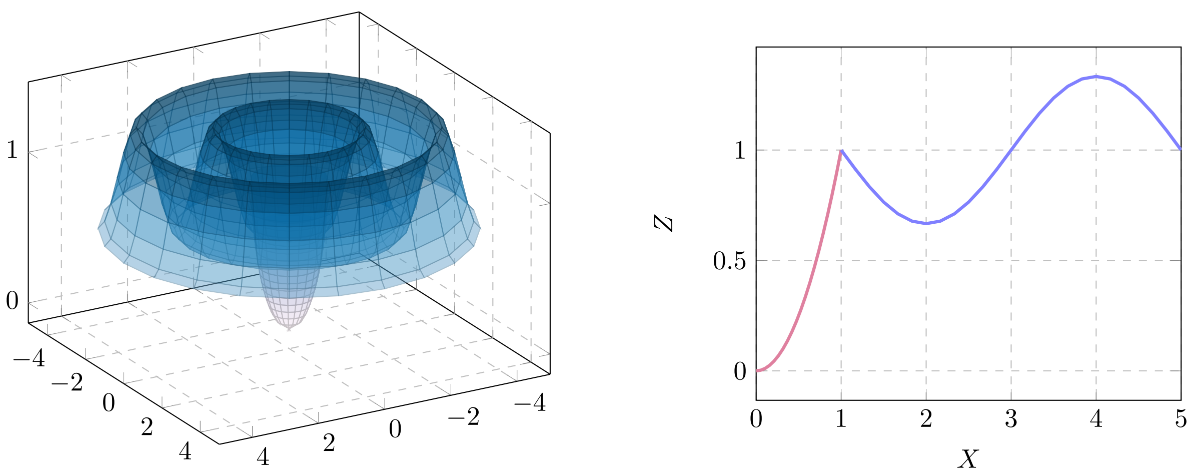


Figura 1.21: Gráfico de función descrita por ramas. A la derecha, corte transversal del plano XZ (con $y = 0$ constante), en el rango $x \in [0, 5]$.

Esta función se comporta como una parábola 2D, hasta alcanzar el borde de la bola unitaria, y luego comienza a oscilar como la función coseno. $\Diamond$

Atención

Para una función vectorial $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ podemos usar dos convenciones de variables:

1. Podemos tomar la variable vectorial $x \in \mathbb{R}^n$, escribiendo $f(x)$.
2. Podemos tomar n variables reales $x_1, \ldots, x_n \in \mathbb{R}$, escribiendo $f(x_1, \ldots, x_n)$.

Durante este curso, ocuparemos ambas, dependiendo del contexto. En el segundo caso, además, cuando $n = 2$ o $n = 3$, escribiremos a veces (x, y) en vez de (x_1, x_2), y (x, y, z) en vez de (x_1, x_2, x_3).

1.5.1 Límites de Funciones

Cuando tomamos una función de una variable $f : [a, b] \to \mathbb{R}$, el límite de la función $\lim_{x\to\bar{x}} f(x)$, con $\bar{x} \in [a, b]$, estaba dado por los límites laterales. Estos representaban "todas las formas" que tenía y de acercarse a x: o bien por la izquierda, o bien por la derecha.
Como ya hemos visto, en $\mathbb{R}^n$ las "formas de acercarse" a un vector fijo pueden ser muy variadas. En efecto, para $\bar{x} \in \mathbb{R}^n$, toda sucesión (x_k) que converja a $\bar{x}$ representa una forma distinta de acercarse a $\bar{x}$. Aún así, para estudiar límites de funciones mantendremos la misma idea que teníamos de límites laterales: Para que el límite $\lim_{x\to\bar{x}} f(x)$ exista, la función f debe converger al mismo valor a través de todas las "formas de acercarse" a $\bar{x}$.

Definición 1.5.4 (Límites de funciones vectoriales)**.** *Sea $f : \Omega \in \mathbb{R}^n \to \mathbb{R}^m$ y sea $\bar{x} \in \mathrm{adh}(\Omega)$. Decimos que un vector $\bar{y} \in \mathbb{R}^m$ es el **límite de f, cuando x tiende a $\bar{x}$** si*

$$\forall (x_k) \subseteq \Omega \ \textit{con} \ x_k \to \bar{x}, \ \textit{se tiene que} \ f(x_k) \to \bar{y}. \tag{1.34}$$

En tal caso, escribimos $\lim\limits_{x\to\bar{x}} f(x) = \bar{y}$.

De la definición anterior, vale la pena destacar algunas observaciones:

1. La definición de límite está dada para todo punto $\bar{x}$ adherente del dominio. Es decir, está definida no sólo para puntos del dominio Ω, donde podemos evaluar f, sino que también para posibles puntos $\bar{x} \in \mathrm{adh}(\Omega) \setminus \Omega$. En estos últimos, no es posible evaluar la función f, pero si aproximar $\bar{x}$ por sucesiones donde f es evaluada.

2. Cuando decimos que $f(x_k) \to \bar{y}$, lo que ocurre implícitamente es que definimos una sucesión $(y_k) \subseteq \mathbb{R}^m$ dada por $y_k = f(x_k)$, y es en esta sucesión donde $\lim\limits_{k\to\infty} y_k = \bar{y}$.

3. Finalmente, para demostrar que un límite no existe (es decir, la negación de (1.34)), necesitamos:

   a) encontrar una sucesión $(x_k) \subseteq \Omega$ con $x_k \to \bar{x}$ tal que el límite $\lim_k f(x_k)$ no exista; o bien,

   b) encontrar dos sucesiones $(x_k^1), (x_k^2) \subseteq \Omega$ convergiendo a $\bar{x}$ tal que $\lim_k f(x_k^1) \neq \lim_k f(x_k^2)$.

   Basta probar una de las dos condiciones anteriores para concluir que el límite $\lim_{x\to\bar{x}} f(x)$ no existe.

Atención

Para demostrar que $\lim_{x \to \bar{x}} f(x) = \bar{y}$, es necesario que $f(x_k) \to \bar{y}$ para toda sucesión (x_k) que converja a $\bar{x}$. Por lo tanto, es necesario trabajar con una sucesión (x_k) con $x_k \to \bar{x}$ arbitraria. Si mostramos que $f(x_k) \to \bar{y}$ para alguna sucesión específica, eso no es suficiente.
Por ejemplo, sea $\Omega = \{(x, y) \in \mathbb{R}^2 \ : \ x > 0, y > 0\}$ y tomemos la función:

$$\begin{aligned} f: \quad \Omega \quad &\to \mathbb{R} \\ (x, y) &\mapsto f(x, y) = x/y \end{aligned}$$

- Si tratamos de calcular $\lim_{(x,y) \to (0,0)} f(x, y)$, podemos considerar la sucesión

$$(x_k, y_k) = (1/k, 1/k) \subseteq \Omega.$$

- Claramente $(x_k, y_k) \to (0, 0)$ y $f(x_k, y_k) = 1$ constante, por lo que $\lim_k f(x_k, y_k) = 1$.

- Podríamos pensar que $\lim_{(x,y) \to (0,0)} f(x, y) = 1$. Sin embargo, esto es incorrecto (de hecho, el límite no existe). La razón, es que solo hemos calculado el límite para una sucesión específica, pero no hemos evaluado **todas las sucesiones posibles**.

Teorema 1.5.5 (Caracterizaciones de límite de funciones)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función, sea $\bar{x} \in \mathrm{adh}(\Omega)$, y sea $\bar{y} \in \mathbb{R}^m$. Las siguientes afirmaciones son equivalentes:*

(a) $\displaystyle\lim_{x \to \bar{x}} f(x) = \bar{y}$.

(b) (Caracterización ε-δ) Para todo $\varepsilon > 0$, existe $\delta > 0$ tal que para todo $x \in \Omega$ se cumple que

$$\|x - \bar{x}\| \leq \delta \implies \|f(x) - \bar{y}\| \leq \varepsilon.$$

(c) (Límite por coordenadas) Para todo $i \in [m]$, $\lim_{x \to \bar{x}} f_i(x) = \bar{y}_i$.

Demostración. Demostraremos el ciclo de implicancias.

$(a) \implies (b)$**:**

- Supongamos que $\displaystyle\lim_{x \to \bar{x}} f(x) = \bar{y}$ según 1.34 (hipótesis).

- Razonemos por contradicción y supongamos que no se tiene (b), es decir,

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x_\delta \in \Omega : \quad \|x_\delta - \bar{x}\| \leq \delta \quad \text{ y } \quad \|f(x_\delta) - \bar{y}\| > \varepsilon.$$

- Construyamos la siguiente sucesión: para cada $k \in \mathbb{N}$ consideremos $\delta = \frac{1}{k}$, entonces:

$$\exists z_k = x_{\frac{1}{k}} \in \Omega : \quad \|z_k - \bar{x}\| \leq \frac{1}{k} \quad \text{ y } \quad \|f(z_k) - \bar{y}\| > \varepsilon.$$

- Como $\|z_k - \bar{x}\| \leq \frac{1}{k} \to 0$, tenemos que $z_k \to \bar{x}$.

- Como $\|f(z_k) - \bar{y}\| > \varepsilon$, tenemos que $f(z_k) \not\to \bar{y}$.

- Por hipótesis: si $z_k \to \bar{x}$, necesariamente $f(z_k) \to \bar{y}$. Lo que es una contradicción con el punto anterior.

- Concluimos que (b) es cierta.

$(b) \implies (c)$**:**

- Queremos demostrar que para todo $i \in [m]$, $\lim_{x \to \bar{x}} f_i(x) = \bar{y}_i$.
  - Fijemos $i \in [m]$ cualquiera.
  - Sea (x_k) una sucesión en Ω con $x_k \to \bar{x}$ cualquiera.
  - Sea $\varepsilon > 0$ cualquiera.
- Usando (b), existe $\delta > 0$ tal que $\|x - \bar{x}\| \leq \delta \implies \|f(x) - \bar{y}\| \leq \varepsilon$.
- En particular, $\|x - \bar{x}\| \leq \delta \implies |f_i(x) - \bar{y}_i| \leq \varepsilon$ (porque $|f_i(x) - \bar{y}_i| \leq \|f(x) - \bar{y}\|$).
- Como $x_k \to \bar{x}$, existe $k_0 \in \mathbb{N}$ tal que $\forall k \geq k_0,\ \|x_k - \bar{x}\| \leq \delta$.
- Juntando los puntos anteriores, concluimos que: $\exists k_0 \in \mathbb{N},\ \forall k \geq k_0,\ |f_i(x) - \bar{y}_i| \leq \varepsilon$.
- Como $\varepsilon > 0$ es arbitrario, hemos demostrado entonces que:

$$\forall \varepsilon > 0, \exists k_0 \in \mathbb{N}, \forall k \geq k_0, |f_i(x_k) - \bar{y}_i| \leq \varepsilon.$$

- Esto significa que $f_i(x_k) \to \bar{y}_i$ (en cuanto a sucesión real).
- Como la sucesión (x_k) es arbitraria, tenemos que $\lim_{x \to \bar{x}} f_i(x) = \bar{y}_i$.
- Como la coordendada $i \in [m]$ es arbitraria, concluimos (c).

$(c) \implies (a)$**:**

- Sea $(x_k) \subseteq \Omega$ una sucesión convergiente a $\bar{x}$ cualquiera.
- Por hipótesis, para todo $i \in [m],\ f_i(x_k) \xrightarrow{k \to \infty} \bar{y}_i$.
- Luego, por definición de convergencia de vectores, tenemos que

$$f(x_k) = (f_1(x_k), \ldots, f_m(x_k)) \xrightarrow{k \to \infty} (\bar{y}_1, \ldots, \bar{y}_m) = \bar{y}.$$

- Como la sucesión (x_k) es arbitraria, concluimos que $\lim_{x \to \bar{x}} f(x) = \bar{y}$.

$\square$

Dado que el límite de funciones está definido a partir de límites de sucesiones, es razonable esperar que

Proposición 1.5.6 (Unicidad del Límite)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función, sea $\bar{x} \in \operatorname{adh}(\Omega)$. Si el límite $\lim_{x \to \bar{x}} f(x)$ existe, entonces es único.*

Demostración.

- Razonando por contradicción, supongamos que existen dos vectores $\bar{y}_1, \bar{y}_2 \in \mathbb{R}^m$ distintos que son límites de f cuando x tiende a $\bar{x}$.
- Como $\bar{x} \in \operatorname{adh}(\Omega)$, existe al menos una sucesión $(x_k) \subseteq \Omega$ tal que $x_k \to \bar{x}$.
- Definamos $(y_k) \subseteq \mathbb{R}^m$ por $y_k = f(x_k)$.
- Por definición de límites de funciones tenemos que $y_k \to \bar{y}_1$ y que $y_k \to \bar{y}_2$.
- Esto es una contradicción con la unicidad del límite de sucesiones (ver Proposición 1.3.5).

□

Proposición 1.5.7 (Álgebra de límites de funciones)**.** *Sean* $f, g : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *dos funciones, sea* $\bar{x} \in \mathrm{adh}(\Omega)$*. Supongamos que el límite* $\lim_{x\to\bar{x}} f(x) = \bar{y}$ *y el límite* $\lim_{x\to\bar{x}} g(x) = \bar{z}$*. Entonces, se tienen las siguientes propiedades:*

1. $\lim\limits_{x\to\bar{x}} (f+g)(x) = \lim\limits_{x\to\bar{x}} f(x) + \lim\limits_{x\to\bar{x}} g(x) = \bar{y} + \bar{z}$.
2. *Para todo* $\alpha \in \mathbb{R}$*, se tiene que* $\lim\limits_{x\to\bar{x}} (\alpha f)(x) = \alpha \lim\limits_{x\to\bar{x}} f(x) = \alpha\bar{y}$.
3. *Si* $m = 1$ *(es decir,* f *y* g *son funciones a valores reales), entonces*
$$\lim_{x\to\bar{x}} (fg)(x) = \Big(\lim_{x\to\bar{x}} f(x)\Big) \cdot \Big(\lim_{x\to\bar{x}} g(x)\Big) = \bar{y} \cdot \bar{z}.$$
4. *Si* $m = 1$ *(es decir,* f *y* g *son funciones a valores reales) y* $\bar{z} \neq 0$*, entonces*
$$\lim_{x\to\bar{x}} \left(\frac{f}{g}\right)(x) = \frac{\lim\limits_{x\to\bar{x}} f(x)}{\lim\limits_{x\to\bar{x}} g(x)} = \frac{\bar{y}}{\bar{z}}.$$

Demostración. Demostraremos las cuatro proposiciones simultáneamente.

- Sea $(x_k) \subseteq \Omega$ una sucesión convergente a $\bar{x}$, y definamos las sucesiones (y_k) y (z_k) dadas por $y_k = f(x_k)$ y $z_k = g(x_k)$, respectivamente.
- Por hipótesis $y_k \to \bar{y}$ y $z_k \to \bar{z}$.
- Usando álgebra de límites de sucesiones (ver Proposiciones 1.2.11 y 1.3.6):
  1. $\lim_k (f+g)(x_k) = \lim_k (y_k + z_k) = \bar{y} + \bar{z}$.
  2. Para todo $\alpha \in \mathbb{R}$, se tiene que $\lim_k (\alpha \cdot f)(x_k) = \lim_k \alpha \cdot y_k = \alpha \cdot \bar{y}$.
  3. Si $m = 1$, entonces $\lim_k (f \cdot g)(x_k) = \lim_k y_k \cdot z_k = \bar{y} \cdot \bar{z}$.
  4. Si $m = 1$ y $\bar{z} \neq 0$, entonces y_k/z_k está bien definida para todo $k \geq k_0$ (con k_0 suficientemente grande) y por lo tanto:
  $$\lim_k \left(\frac{f}{g}\right)(x_k) = \lim_k \frac{y_k}{z_k} = \frac{\bar{y}}{\bar{z}}.$$
- Como (x_k) era una sucesión convergente cualquiera, concluimos que los resultados son ciertos para los límites de funciones.

□

Así como algunos resultados de álgebra solo funcionan para el caso $m = 1$, lo mismo pasa con el Teorema del Sándwich.

Proposición 1.5.8 (Teorema del Sándwich - límites de funciones)**.** *Sean* $f, g, h : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ *tres funciones, sea* $\bar{x} \in \mathrm{adh}(\Omega)$*. Supongamos que*

(i) *Existe* $\delta > 0$ *tal que para todo* $x \in B(x, \delta) \cap \Omega$*, se tiene que* $f(x) \leq g(x) \leq h(x)$.

(ii) $\lim\limits_{x\to\bar{x}} f(x) = \lim\limits_{x\to\bar{x}} h(x) = \lambda$.

Entonces, $\lim_{x\to\bar{x}} g(x)$ *existe y es igual a* λ*.*

Demostración. Notemos que en esta demostración debemos usar las hipótesis (i) y (ii).

- Sean
  - $(x_k) \subseteq \Omega$ una sucesión con $x_k \to \bar{x}$ cualquiera.
  - $\delta > 0$ de la hipótesis (i).
  - Las sucesiones reales $(f(x_k))_{k\in\mathbb{N}}$, $(g(x_k))_{k\in\mathbb{N}}$ y $(h(x_k))_{k\in\mathbb{N}}$.
- Por Teorema 1.3.4 existe $k_0 \in \mathbb{N}$ tal que: $\forall k \geq k_0,\ \|x_k - \bar{x}\| \leq \delta$.
- Es decir, $\forall k \geq k_0,\ x_k \in B(\bar{x}, \delta) \cap \Omega$.
- Por hipótesis (i): $\forall k \geq k_0, f(x_k) \leq g(x_k) \leq h(x_k)$.
- Por hipótesis (ii), usando Teorema del Sándwich de sucesiones reales (Proposición 1.2.13), concluimos que $\lim_k g(x_k) = \lambda$.
- Como la sucesión es arbitraria, tenemos que $\lim_{x\to\bar{x}} g(x) = \lambda$.

□

El siguiente teorema nos permitirá simplificar el estudio de límite de funciones, cuando las funciones involucradas están definidas por ramas. Al definir una función $f : \Omega \to \mathbb{R}^m$ por ramas, lo que estamos haciendo es particionar el dominio Ω en varios pedazos $A_1, A_2, \ldots, A_p$, donde para cada partición A_j, f toma una forma distinta. Cuando estudiamos un límite de f, tenemos que estudiar las sucesiones (x_k) en cada una de las particiones pero, en principio, también tenemos que considerar las sucesiones que alternan entre particiones. El siguiente teorema nos dice que estas sucesiones alternantes pueden ser obviadas.

Teorema 1.5.9 (Límite por finitas partes). *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función, sea $\bar{x} \in \mathrm{adh}(\Omega)$, y sea $\bar{y} \in \mathbb{R}^m$. Sea $\{A_j \ :\ j = 1, \ldots, p\}$ una colección finita de conjuntos no vacíos tal que $\Omega = \bigcup_{j=1}^{p} A_j$. Entonces,*

$$\lim_{x\to\bar{x}} f(x) = \bar{y} \iff \forall j \in [p],\ \forall (x_k) \subseteq A_j \text{ con } x_k \to \bar{x},\ \text{se tiene que } f(x_k) \to \bar{y}. \tag{1.35}$$

Demostración. Demostremos cada implicancia por separado.

$(\Longrightarrow)$

- Supongamos que $\lim_{x\to\bar{x}} f(x) = \bar{y}$.
- Por definición (ver 1.5.4): $\forall (x_k) \subseteq \Omega$ con $x_k \to \bar{x},$ se tiene que $f(x_k) \to \bar{y}$.
- Reduciendo al caso particular de las sucesiones contenidas en cada partición, tenemos que:

$$\forall j \in [p],\ \forall (x_k) \subseteq A_j \text{ con } x_k \to x,\ \text{se tiene que } f(x_k) \to \bar{y}.$$

$(\Longleftarrow)$

- Supongamos que $\forall j \in [p],\ \forall (x_k) \subseteq A_j$ con $x_k \to x,$ se tiene que $f(x_k) \to \bar{y}$ (hipótesis).
- Para cada partición $j \in [p]$, podemos considerar la función f restringida a esa partición, es decir:

$$\begin{aligned} f^j : A_j &\to \mathbb{R}^m \\ x &\mapsto f^j(x) = f(x). \end{aligned}$$

- Con las funciones restringidas, la hipótesis se traduce a que $\lim_{x\to\bar{x}} f^j(x) = \bar{y}$, para todo j tal que $\bar{x} \in \mathrm{adh}(A_j)$ (si no está en la adherencia no se puede tomar el límite por definición).

- Llamemos $J_{\bar{x}} = \{j \in [p] : \ \bar{x} \in \text{adh}(A_j)\}$ los índices donde se puede calcular el límite.
- Queremos demostrar que $\lim_{x \to \bar{x}} f(x) = \bar{y}$. Vamos a usar la caracterización $\varepsilon - \delta$ dada por el Teorema 1.5.5. Sea $\varepsilon > 0.$
- Para $j \notin J_{\bar{x}}$, usando la Proposición 1.4.7, tenemos que existe $\delta_j > 0$ tal que $B(\bar{x}, \delta_j) \cap A_j = \emptyset$.
- Para $j \in J_{\bar{x}}$, usando el Teorema 1.5.5 en las funciones restringidas, tenemos que:

$$\exists \delta_j > 0, \ \forall x \in B(\bar{x}, \delta_j) \cap A_j, \|f^j(x) - \bar{y}\| \leq \varepsilon. \tag{1.36}$$

- Tomemos $\delta_0 = \min\{\delta_j \ : \ j \in [p]\} > 0$.
- Sea $x \in \Omega$ cualquiera tal que $\|x - \bar{x}\| \leq \delta_0$.
- Como $\Omega = \bigcup_{j=1}^{p} A_j$, entonces existe $j^* \in [p]$ tal que $x \in A_{j^*}$.
- Necesariamente $j^* \in J_{\bar{x}}$ porque $B(\bar{x}, \delta_0) \cap A_j = \emptyset$ para los otros índices.
- Por (1.36) y recordando que $f^j(x) = f(x)$, tenemos que: $\|f(x) - \bar{y}\| \leq \varepsilon$.
- Como $x \in \Omega$ era arbitrario, concluimos que

$$\forall x \in \Omega : \quad \|x - \bar{x}\| \leq \delta_0 \implies \|f(x) - \bar{y}\| \leq \varepsilon.$$

- Como ε era arbitrario, tenemos la caracterización $\varepsilon - \delta$:

$$\forall \varepsilon > 0, \ \exists \delta_0 > 0, \ \forall x \in \Omega : \quad \|x - \bar{x}\| \leq \delta_0 \implies \|f(x) - \bar{y}\| \leq \varepsilon,$$

  que nos permite concluir que $\lim_{x \to \bar{x}} f(x) = \bar{y}$.

$\square$

Nota: Límite por partes y límites laterales

Existe un paralelo natural entre el límite por finitas partes en funciones en varias variables y los límites laterales en las funciones reales. En el caso de $\mathbb{R}^n$, no se puede reducir a "acercarse por un lado o por el otro", porque hay muchas direcciones desde donde "acercarse". Pero sí podemos separar el espacio y entonces elegir caminos representativos de todo el espacio. Ambas estrategias nos permiten reducir el análisis necesario para poder concluir la existencia de un límite.

Por ejemplo, en la Figura 1.22 tenemos que:

- $x \in A_1, \quad x \in \text{adh}(A_1), \ x \in \text{adh}(A_2), \ x \notin \text{adh}(A_3)$.
- $y \in A_3, \quad y \in \text{adh}(A_1), \ y \in \text{adh}(A_2), \ y \in \text{adh}(A_3)$.
- $z \in A_3, \quad z \notin \text{adh}(A_1), \ z \notin \text{adh}(A_2), \ z \in \text{adh}(A_3)$.

Por lo que para calcular un límite en x, basta considerar sucesiones en A_1 y A_2 por separado. Esos dos "caminos" representan todas las formas de acercarse a x.
De forma similar, para calcular un límite en y se necesitan considerar sucesiones por A_1, A_2 y A_3 por separado; y para z solo las sucesiones de A_3.

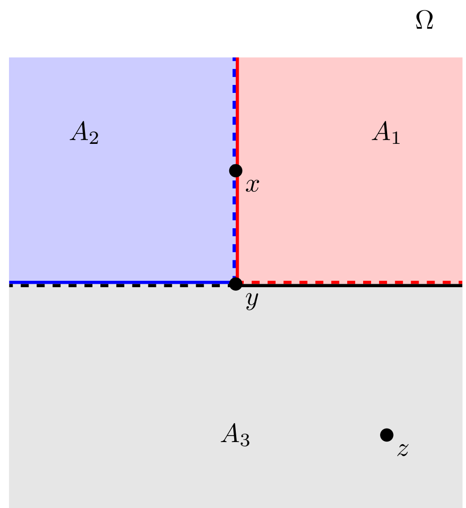


Figura 1.22: Partición de un espacio Ω en A_1, A_2 y A_3.

1.5.2 Continuidad de funciones

De la noción de límites de funciones de varias variables, podemos extender naturalmente la noción de continuidad.

Definición 1.5.10. *Sea $f : \Omega \to \mathbb{R}^m$ y sea $\bar{x} \in \Omega$. Decimos que f es* ***continua en $\bar{x}$*** *si*

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

Además, decimos que

- *f es continua en un conjunto $A \subseteq \Omega$, si lo es para todo punto de A.*
- *f es continua (a secas), si es continua en todo su dominio Ω.*

Como veremos a lo largo del curso, la continuidad de funciones es una propiedad fundamental que sirve como base para otras construcciones, como derivadas e integrales. Las funciones continuas nos permiten asegurar que los valores de una función son estables con respecto a tomar límites. Esta estabilidad nos permite admitir errores numéricos con la tranquilidad que si el error es pequeño, los valores de la función no pueden variar mucho. También nos permiten construir procesos iterativos basados en hecho que, si continuamos infinitamente, la tendencia que observamos en los datos generados de hecho se tiene que parecer al valor de la función en el punto de convergencia.

Atención

La continuidad solo se puede estudiar en puntos del dominio Ω, mientras que los límites se estudian para puntos adherentes en $\text{adh}(\Omega)$.
Para $f : \Omega \to \mathbb{R}^m$ y $\bar{x} \in \text{adh}(\Omega) \setminus \Omega$ tal que $\lim_{x \to \bar{x}} f(x) = \bar{y}$, no podemos decir que $f(\bar{x}) = \bar{y}$.
Simplemente, f no puede evaluarse en puntos fuera de Ω.

Una consecuencia directa de la definición de continuidad, es que las funciones continuas heredan la misma álgebra que se tiene para límites de funciones.

Proposición 1.5.11 (Álgebra de funciones continuas)**.** *Sean $f, g : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ dos funciones y sea $\bar{x} \in \Omega$. Asumamos que f y g son continuas en $\bar{x}$. Entonces, se tiene que*

1. $f + g$ *es continua en* $\bar{x}$*.*

2. Para todo $\alpha \in \mathbb{R}$*,* αf *es continua en* $\bar{x}$*.*

3. Si $m = 1$*,* fg *es continua en* $\bar{x}$*.*

4. Si $m = 1$ *y* $g(\bar{x}) \neq 0$*, entonces la función* $h = \frac{f}{g}$ *está bien definida en* $B(\bar{x}, \delta) \cap \Omega$ *para* $\delta > 0$ *lo suficientemente pequeño, y además* h *es continua en* $\bar{x}$*.*

Demostración. Las propiedades 1,2 y 3 son consecuencia directa del álgebra de límites de funciones (ver Proposición 1.5.7). Demostremos la propiedad 4.

- Definamos $r = |g(\bar{x})| > 0$.
- Usando la caracterización $\varepsilon - \delta$ (ver Teorema 1.5.5) para $\varepsilon = r/2$, tenemos que existe $\delta > 0$ tal que
$$\forall x \in B(\bar{x}, \delta) \cap \Omega, |g(x) - g(\bar{x})| \leq \frac{r}{2}.$$
- Por lo tanto, para todo $x \in B(\bar{x}, \delta) \cap \Omega$ tenemos que
$$\begin{aligned} r - |g(x)| &= |g(\bar{x})| - |g(x)| \\ &= |g(\bar{x}) - g(x) + g(x)| - |g(x)| \\ &\leq |g(\bar{x}) - g(x)| + |g(x)| - |g(x)| \\ &= |g(\bar{x}) - g(x)| \\ &\leq \frac{r}{2}. \end{aligned}$$
- Reorganizando, concluimos que para todo $x \in B(\bar{x}, \delta) \cap \Omega :\ |g(x)| \geq \frac{r}{2} > 0$.
- Concluimos que h está bien definida en $B(\bar{x}, \delta) \cap \Omega$. Luego, la continuidad de h es una aplicación del límite de cociente de funciones (ver Proposición 1.5.7).

□

Para finalizar este capítulo, incluiremos un último criterio de continuidad que corresponde a la composición de funciones continuas. Junto con el álgebra de funciones continuas, y la continuidad de funciones de una variable, podremos construir muchas funciones continuas de manera casi automática.

Nota: Rango de funciones

Para una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$, denotamos su imagen o rango como $f(\Omega)$. Es decir, $f(\Omega)$ es el conjunto dado por
$$f(\Omega) = \{y \in \mathbb{R}^m \ :\ \exists x \in \Omega, y = f(x)\}.$$

Proposición 1.5.12 (Composición de funciones continuas)**.** *Sea* $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *y sea* $g : \Lambda \subseteq \mathbb{R}^m \to \mathbb{R}^p$*. Sea* $\bar{x} \in \Omega$ *y supongamos que*

(i) $f(\Omega) \subseteq \Lambda$*.*

(ii) f *es continua en* $\bar{x}$*.*

(iii) g *es continua en* $\bar{y} = f(\bar{x})$*.*

Entonces, la composición $g \circ f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^p$ *es continua en* $\bar{x}$*.*

Demostración. Queremos demostrar que $\lim_{x\to\bar{x}}(g\circ f)(x) = (g\circ f)(\bar{x})$.

- Sea $(x_k)\subseteq\Omega$ una sucesión cualquiera con $x_k\to\bar{x}$.
- Como $f(\Omega)\subseteq\Lambda$, podemos considerar la sucesión $(y_k)\subseteq\Lambda$ definida por $y_k = f(x_k)$.
- Como f es continua en $\bar{x}$, tenemos que $y_k\to\bar{y} = f(\bar{x})$.
- Como g es continua en $\bar{y}$, tenemos que $g(y_k)\to g(\bar{y})$.
- Podemos concluir que:

$$(g\circ f)(x_k) = g(f(x_k)) = g(y_k)\to g(\bar{y}) = g(f(\bar{x})) = (g\circ f)(\bar{x}).$$

- Como la sucesión (x_k) es arbitraria, concluimos que $\lim_{x\to\bar{x}}(g\circ f)(x) = (g\circ f)(\bar{x})$.

□

1.6 Selección de Problemas

A continuación, se presenta una selección de problemas confeccionada por los autores, para ejercitar los contenidos hasta aquí expuestos. Se solicita al lector o lectora, considerar la siguiente simbología al momento de trabajar en ellos: problemas destacados de la forma $(\star)$ son representativos de controles de cátedra, y destacados de la forma $(\star\star)$ indican que son problemas de dificultad elevada.

1.6.1 El Espacio $\mathbb{R}^n$

1. $(\star)$ Considere en $\mathbb{R}^n$ el producto interno $\langle\cdot,\cdot\rangle$ y la norma euclidiana $\|\cdot\|$.

   a) Muestre que la norma euclidiana cumple la siguiente identidad,

   $$(\forall x, y\in\mathbb{R}^n)\quad \|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

   b) Sean $v_1, v_2, \ldots, v_m\in\mathbb{R}^n$ vectores ortogonales entre sí. Demuestre que

   $$\|v_1+v_2+\ldots+v_m\|^2 = \|v_1\|^2 + \|v_2\|^2 + \ldots + \|v_m\|^2.$$

2. $(\star)$ Muestre que la norma euclidiana $\|\cdot\|$ en $\mathbb{R}^n$ cumple la llamada *identidad de polarización*

   $$\forall x, y\in\mathbb{R}^n\quad \langle x, y\rangle = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right).$$

3. $(\star\star)$ Además de la norma euclidiana definida en el curso, en un contexto más general, en un espacio vectorial cualquiera, toda función que cumpla las propiedades de la Proposición 1.1.4 (Positividad, Separación, Homogeneidad y Desigualdad Triangular) recibe el nombre de *norma*.

   Muestre que en $\mathbb{R}^n$, la *norma infinito*

   $$\|x\|_\infty = \max_{k=1,\ldots,n} |x_k|$$

   se trata, efectivamente, de una norma.

4. Muestre que si cambia la norma euclidiana por la norma infinito, la identidad demostrada en 1.a) ya no es cierta. Para ello, encuentre un par de vectores adecuados para que la igualdad no se cumpla.

5. ($\star\star$) Definimos $\mathcal{C}([0,1],\mathbb{R})$ como el espacio vectorial de las funciones continuas definidas desde el intervalo cerrado $[0,1]$ a valores reales; dotado de la suma y la ponderación por escalar. Muestre que en este espacio, la función

$$\begin{aligned} \|\cdot\|_1 \ : \ \mathcal{C}([0,1],\mathbb{R}) &\to \mathbb{R} \\ f &\mapsto \|f\|_1 = \int_0^1 |f(t)|dt \end{aligned}$$

se trata de una norma.

6. En mecánica clásica, se dice que una fuerza realiza trabajo cuando hay un desplazamiento de su punto de aplicación. En este mismo contexto, el trabajo de la fuerza sobre ese cuerpo será equivalente a la energía necesaria para poder desplazarlo en una cierta dirección; y en términos matemáticos, si $\vec{F}$ es la fuerza aplicada sobre el cuerpo y $\vec{d}$ es su desplazamiento, se puede definir como

$$W = \langle \vec{F}, \vec{d} \rangle.$$

Considere en este problema un objeto que se desea mover en una línea recta, desde el origen de coordenadas hasta el punto $\vec{d} = (5,4,3)$ en tres dimensiones, y las fuerzas representadas por los vectores detallados a continuación

$$\vec{F_1} = (1,1,1)\ , \quad \vec{F_2} = (0,1,-2)\ , \quad \vec{F_3} = (3,-1,2)\ .$$

a) ¿Cuál de las tres fuerzas involucradas genera más trabajo para mover el objeto en cuestión?

b) ¿Cómo cambian sus resultados, si ahora se busca el desplazamiento $\vec{d} = (-1,1,-1)$?

c) ¿Qué ocurre si la fuerza aplicada es perpendicular al desplazamiento buscado?
Hint: Recuerde que $\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\|\|\vec{y}\|\cos(\theta)$, donde θ es el ángulo subtendido por los vectores $\vec{x}$ e $\vec{y}$.

d) A partir de sus observaciones hasta aquí, ¿cómo debe ser el vector de fuerzas usado para que genere un mayor trabajo en el desplazamiento buscado?

1.6.2 Sucesiones Reales y Vectoriales

1. ($\star$) En matemáticas, una de las aplicaciones de las sucesiones que hemos estudiado en clases, corresponde al concepto de serie numérica, que en términos intuitivos es una suma infinita de la forma

$$a_1 + a_2 + a_3 + \ldots + a_k + a_{k+1} + \ldots = \sum_{k=1}^{\infty} a_k,$$

donde cada $a_k \in \mathbb{R}$ se puede interpretar como el término general de una sucesión. Una forma de trabajar con esta intuición y poder formalizarla en términos matemáticos, será la de definir el concepto de "suma parcial" de una serie

$$s_n = \sum_{k=1}^{n} a_k,$$

e intentar determinar el valor del límite

$$L = \lim_{n\to\infty} s_n.$$

El criterio de convergencia de series que estudiaremos será el siguiente: *Sea $(a_k)_{k\in\mathbb{N}}$ una sucesión de términos positivos, y supongamos que el siguiente límite existe*

$$r = \lim_{k\to\infty} \sqrt[k]{a_k}.$$

Dependiendo del valor de r, tenemos los siguientes escenarios:

- *Si $r < 1$, entonces la serie $\sum_{k=1}^{\infty} a_k$ converge.*
- *Si $r > 1$, entonces la serie $\sum_{k=1}^{\infty} a_k$ diverge.*
- *Si $r = 1$, entonces el criterio no es concluyente (no sabemos qué pasa con la serie).*

a) Muestre que para cualquier $x \in \mathbb{R}$, la siguiente serie numérica converge

$$\sum_{k=1}^{\infty} \frac{x^n}{n!}.$$

b) Considere ahora la serie numérica dada por

$$\sum_{k=1}^{\infty} \frac{(k!)^2 x^k}{(2k)!}$$

Muestre que esta serie converge si $x < 4$, pero diverge si $x > 4$.

2. Estudie la convergencia de las siguientes sucesiones:

   a) $A_n = \left(\sin(n\pi),\ \cos(n\pi),\ \dfrac{1 - \cos(n\pi)}{n^n} \right).$

   b) $B_n = \left(\dfrac{e^{\frac{1}{2n}} - 1}{2n},\ (2n+1)\sin\left(\dfrac{1}{3n}\right) \right).$

3. ($\star$) Indique si la siguiente afirmación es verdadera o falsa, justificando adecuadamente su respuesta: La sucesión

$$x_n = \left(\frac{\cos(e^n)}{n}, \frac{2n+4}{n-3}, -2 + \frac{1}{n!} \right)$$

   es convergente en $\mathbb{R}^3$.

4. ($\star\star$) Estudie la convergencia de la sucesión en $\mathbb{R}^2$ dada por

$$x_n = \left(\frac{a_n b_n}{|a_n| + b_n^2}, \frac{b_n(a_n^2 + b_n^2)^{3/2}}{(a_n^2 + b_n^2)^2 + b_n^2} \right),$$

   donde $(a_n, b_n) \to (0, 0)$.

5. ($\star\star$) Sea $(x_n, y_n)_{n\in\mathbb{N}}$ una sucesión en $\mathbb{R}^2$ tal que $\lim\limits_{n\to\infty} x_n = 0$. Decida la veracidad de las siguientes afirmaciones, mostrando por qué son verdaderas o exhibiendo un contraejemplo.

   a) $(x_n, y_n)_{n\in\mathbb{N}}$ es convergente si y sólo si $\lim_{n\to\infty} y_n = 0$.

   b) $(x_n, y_n)_{n\in\mathbb{N}}$ es convergente si y sólo si $(y_n)_{n\in\mathbb{N}}$ es convergente.

   c) Si $(y_n)_{n\in\mathbb{N}}$ es acotada, entonces $(x_n, y_n)_{n\in\mathbb{N}}$ es convergente.

1.6.3 Conjuntos Abiertos y Cerrados

1. Sean $A, B \subseteq \mathbb{R}^n$ dos conjuntos no vacíos. Demuestre las siguientes propiedades.

   a) $\text{adh}(A \cap B) \subseteq \text{adh}(A) \cap \text{adh}(B)$.

   b) $\text{adh}(A \cup B) = \text{adh}(A) \cup \text{adh}(B)$.

   c) $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$.

   d) $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$.

   En particular, en los casos en los que se tiene la inclusión, encuentre ejemplos que expliciten que en general, la igualdad no es cierta.

2. ($\star$) Se definen los conjuntos $A, B \subseteq \mathbb{R}^2$ dados por

$$A = \{(x,y) \in \mathbb{R}^2 \ : \ x \leq y < 3 - x^2\}\ , \quad B = \{(x,y) \in \mathbb{R}^2 \ : \ 1 \leq x \leq 7,\ y \geq 3\}.$$

   a) Bosqueje en el plano cartesiano ambos conjuntos.

   b) Determine interior y adherencia de ambos conjuntos.

   c) Usando la parte anterior, indique si alguno de los conjuntos en cuestión es abierto o cerrado. Justifique su respuesta.

   d) Se dice que un conjunto $C \subseteq \mathbb{R}^2$ es *acotado*, si existe una bola $B((x_0, y_0), r)$ tal que $C \subseteq B((x_0, y_0), r)$. Determine si A y B son o no conjuntos acotados. Justifique sus respuestas.

3. ($\star$) Considere el conjunto $C \subseteq \mathbb{R}^2$ dado por

$$C = \{(x,y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leq 1,\ |x| < |y|\}.$$

   Bosqueje el conjunto C, e identifique su interior, adherencia y frontera.

1.6.4 Límites y Continuidad

1. Determine si la siguiente afirmación es verdadera o falsa, justificando su respuesta en cualquier caso: Si $f, g, h : \mathbb{R}^n \to \mathbb{R}$ son tres funciones tales que

$$f(x) = g(x) + h(x).$$

   Si existe el límite de f cuando $x \to a$, entonces los límites de g y h cuando $x \to a$ también existen, y cumplen que

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) + \lim_{x \to a} h(x).$$

2. ($\star$) Calcule (si es que existen) los siguientes límites, justificando sus respuestas según corresponda.

   a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2+y^2}$.

   b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{y}{\sqrt{x^2+y^2}}$.

   c) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{(x^2-y)^2}{x^2+y^2}$.

   d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{e^{x+y}-1}{1-\cos(x+y)}$.

3. ($\star\star$) Estudie la existencia de los siguientes límites, calculando aquellos que existan:

   a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\text{sen}(x)+\text{sen}(y)}{x+y}$.

   b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\text{sen}(x)-\text{sen}(y)}{x+y}$.

Indicación: *Busque alguna identidad trigonométrica sobre la suma o resta de senos.*

4. ($\star$) Estudie la continuidad de las siguientes funciones, distinguiendo los casos $(x,y) \neq (0,0)$ y $(x,y) = (0,0)$.

   a)
$$f(x,y) = \begin{cases} \dfrac{x^2y^3}{x^4+y^6} & \text{si } (x,y) \neq (0,0) \\ \\ 0 & \text{si } (x,y) = (0,0) \end{cases}.$$

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

5. ($\star\star$) Considere la función dada por
$$f(x,y) = \begin{cases} \dfrac{\ln(1+x^2+y^2)+x^3+y^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ \alpha & \text{si } (x,y) = (0,0) \end{cases}$$

   Determine el valor de α que hace que f sea una función continua.

6. ($\star$) Repita la tarea recién realizada, para encontrar ahora el valor de β que hace que g sea continua.
$$g(x,y) = \begin{cases} \dfrac{1-\cos\left(\sqrt{x^2+y^2}\right)}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ \beta & \text{si } (x,y) = (0,0) \end{cases}.$$

1.6.5 Misceláneos

1. Pruebe que el siguiente límite no existe
$$\lim_{(x,y)\to(0,0)} \frac{x^4+y^4}{x^3+y^3}$$

   Para ello considere las sucesiones $(\frac{1}{n}, \frac{1-n}{n^2})$ y $(\frac{1}{n}, 0)$, las cuales convergen a $(0,0)$.

2. ($\star$) Indique si las siguientes afirmaciones son verdaderas o falsas. Justifique todas sus respuestas, ya sea con ejemplos, contraejemplos o demostraciones según corresponda.

   a) La sucesión $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^3$ dada por $x_n = \left(\frac{\cos(n!)}{n^2}, \frac{2n-4n^2}{n^2+10}, \frac{1}{\ln(n+2)}\right)$ no es convergente.

   b) Si $A \subseteq \mathbb{R}^n$ es un conjunto abierto y cerrado al mismo tiempo, entonces $\mathrm{Fr}(A) = \varnothing$.

   c) Si $f, g : \mathbb{R}^2 \to \mathbb{R}^4$ son dos funciones tales que $f+g$ es continua en $x_0 \in \mathbb{R}^3$, entonces f y g son continuas en x_0 también.

   d) Si $x, y \in \mathbb{R}^n$ son dos vectores tales que
$$\|x+y\| = \|x\| + \|y\|,$$
   entonces al menos unos de los vectores es cero.

CAPÍTULO 2

Diferenciabilidad en $\mathbb{R}^n$

Tal como hicimos con los límites y la continuidad de las funciones de varias variables al final del Capítulo 1, en la presente sección revisaremos cómo extender la noción de diferenciabilidad en varias variables, comenzando con funciones cuya imagen está en $\mathbb{R}$ para después ver el caso en el que las funciones llegan a $\mathbb{R}^m$.

Para motivar nuestro estudio, recordemos la interpretación que tiene la derivada en una variable, y analicemos gráficamente las diferencias que encontraremos en varias variables.

2.1 Derivadas Direccionales y Parciales

Ya sabemos, a partir de cursos anteriores, que la derivada de una función nos da información con respecto a su crecimiento. Sólo para poder hacer la comparación con el caso multivariable, enunciemos su definición para el caso unidimensional.

Definición 2.1.1 (Derivada en $\mathbb{R}$)**.** *Sea $f : (a, b) \to \mathbb{R}$ una función, y $x_0 \in (a, b)$. Diremos que f es una función* derivable *en x_0, si el siguiente límite existe*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{2.1}$$

En tal caso, denotamos al límite $f'(x_0)$ y lo llamamos derivada de f en x_0.

Nota:

Haciendo el cambio de variables $h = x - x_0$, podemos reescribir el límite que define a la derivada como

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Cuando $f'(x_0)$ existe, es posible entonces afirmar que

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h} = 0 \iff \lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - f'(x_0)h|}{|h|} = 0. \tag{2.2}$$

La ecuación (2.2), es la que más adelante permitirá extender a varias variables este concepto.

La interpretación geométrica de estos cálculos, corresponde a la siguiente: $f'(x_0)$ entrega como resultado *la pendiente de la recta tangente a la gráfica de $f(x)$, en el punto $x = x_0$.*

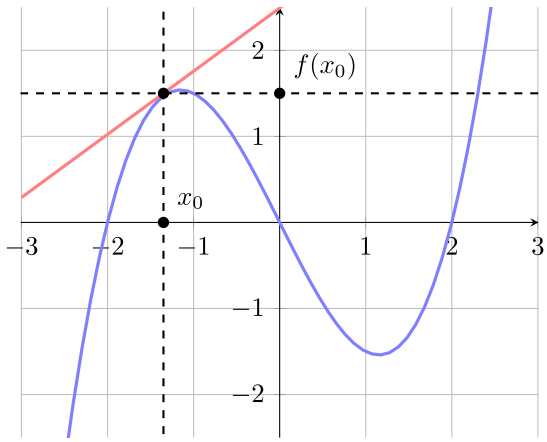


Figura 2.1: Interpretación gráfica de $f'(x_0)$, como la pendiente de la recta tangente a f en x_0.

En el caso de estudiar una función de varias variables, en primera instancia no es tan claro cómo extender esta noción. Basta con tomar una función como

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

y notar que, dependiendo de cuál sea la variable perturbada por el límite, se obtienen diferentes resultados. No es muy difícil calcular, por ejemplo, que se tiene

$$\begin{aligned}
\lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} &= \lim_{h\to 0} \frac{(x_0 + h)^2 - y_0^2 - (x_0^2 - y_0^2)}{h} \\
&= \lim_{h\to 0} \frac{\cancel{x_0^2} + 2x_0 h + h^2 - \cancel{y_0^2} - \cancel{x_0^2} + \cancel{y_0^2}}{h} \\
&= \lim_{h\to 0} \frac{\cancel{h}(2x_0 + h)}{\cancel{h}} \\
&= 2x_0,
\end{aligned}$$

y de manera análoga, se puede calcular también

$$\lim_{h\to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} = -2y_0$$

Más aún, si perturbamos ambas coordenadas de forma simultánea, se puede calcular

$$\lim_{h\to 0} \frac{f(x_0 + h, y_0 + h) - f(x_0, y_0)}{h} = 2x_0 - 2y_0,$$

y en todos los casos, dependiendo del valor específico del par ordenado $(x_0, y_0) \in \mathbb{R}^2$, es posible obtener resultados positivos o negativos, para un mismo punto en el que estamos calculando "una especie" de derivada. Esto dice relación con la coordenada que estamos usando para calcular, como se puede apreciar en la Figura (2.2)

2.1.1 La Derivada Direccional

En resumidas cuentas, podemos concluir que en el caso de una función de varias variables, lo que observemos como "derivada", dependerá fuertemente de la *dirección* que usemos en el espacio vectorial para "mirar". Esta situación, es la que inspira la siguiente definición.

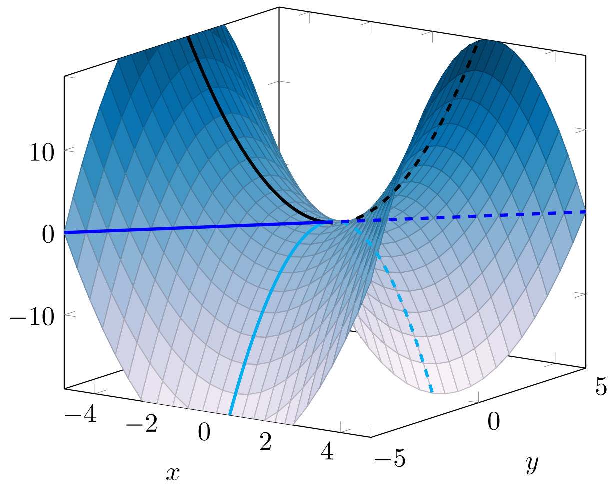


Figura 2.2: Comportamiento de $f(x,y) = x^2 - y^2$ perturbando en eje x (línea nega), eje y (línea celeste) y ambas (línea azul). Se aprecia comportamiento convexo, cóncavo y constante, dependiendo del eje utilizado.

Definición 2.1.2 (Derivada Direccional)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, con Ω un conjunto abierto. Diremos que f tiene derivada direccional en $x_0 \in \Omega$, según la dirección (vector) $v \in \mathbb{R}^n$, si el siguiente límite existe*

$$f'(x_0; v) = \lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h}.$$

Nota: Relación con la derivada unidimensional

Si para cada $x_0 \in \Omega$ y $v \in \mathbb{R}^n$ específicos definimos la función

$$\varphi(t) = f(x_0 + tv),$$

podemos identificar en la definición de derivada direccional lo siguiente:

$$f'(x_0; v) = \lim_{h \to 0} \frac{\varphi(h) - \varphi(0)}{h} = \varphi'(0),$$

que corresponde finalmente a una derivada en $\mathbb{R}$. Por ende, naturalmente, la derivada direccional heredará todas las propiedades que ya conocemos de la derivada unidimensional.

Ejemplo 2.1.3. Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = \begin{cases} \frac{x^3 y \operatorname{sen}(xy^2)}{x^6 + y^6} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}.$$

Si deseamos calcular la derivada direccional de f en el origen, según una dirección $(u,v) \in \mathbb{R}^2$ cualquiera, debemos calcular como se desarrolla a continuación.

$$
\begin{aligned}
f'((0,0);(u,v)) &= \lim_{h\to 0} \frac{f((0,0)+h(u,v)) - f(0,0)}{h}\\
&= \lim_{h\to 0} \frac{f(hu,hv) - f(0,0)}{h}\\
&= \lim_{h\to 0} \frac{\frac{h^4u^3v\operatorname{sen}(h^3uv^2)}{h^6(u^6+v^6)} - 0}{h}\\
&= \lim_{h\to 0} \frac{h^4u^3v\operatorname{sen}(h^3uv^2)}{h^7(u^6+v^6)}\\
&= \lim_{h\to 0} \frac{u^3v\operatorname{sen}(h^3uv^2)}{h^3(u^6+v^6)}\\
&= \frac{u^3v}{u^6+v^6}\cdot \lim_{h\to 0} \frac{\operatorname{sen}(h^3uv^2)}{h^3}\cdot\frac{uv^2}{uv^2}\\
&= \frac{u^4v^3}{u^6+v^6}\cdot \lim_{h\to 0} \frac{\operatorname{sen}(h^3uv^2)}{h^3uv^2}\\
&= \frac{u^4v^3}{u^6+v^6},
\end{aligned}
$$

donde hemos usado el límite conocido

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

$\Diamond$

2.1.2 La Derivada Parcial

Tal como suele pasarnos al estudiar el conjunto $\mathbb{R}^n$, direcciones que serán particularmente importantes de estudiar corresponderán a aquellas dadas por los vectores canónicos, $\{e_j\}_{j=1}^n = \{e_1, e_2, \ldots, e_n\}$, donde

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \longrightarrow \text{j-ésima coordenada}$$

Por ende, las derivadas direccionales en *estas* direcciones, tendrá un nombre especial, y serán estudiadas con más profundidad.

Definición 2.1.4 (Derivada Parcial)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función, con Ω un conjunto abierto. Llamaremos **derivada parcial** de f con respecto a x_j en $x_0 \in \Omega$, a la expresión*

$$\frac{\partial f}{\partial x_j}(x_0) = \lim_{h\to 0} \frac{f(x_0 + he_j) - f(x_0)}{h}.$$

Atención

Las derivadas parciales corresponden a perturbar sólo una de las coordenadas del punto en el que estamos derivando. Además, como una derivada parcial en un punto $x_0 \in \Omega$ específico es en el fondo una derivada de $\mathbb{R}$ en $\mathbb{R}$, a través de la función

$$\varphi_j(t) = f(x_0 + te_j)$$

y recordando que

$$\frac{\partial f}{\partial x_j}(x_0) = f'(x_0; e_j) = \varphi_j'(0),$$

podemos deducir directamente que las derivadas parciales cumplen **todas las reglas de cálculo** que ya conocemos para obtener derivadas en $\mathbb{R}$. En particular, se cumplen las leyes de la suma, de la ponderación por escalar, del producto, del cuociente y de la cadena.

Ejemplo 2.1.5. Para la función

$$g(x,y) = \frac{xy}{x^2+y^2+1},$$

sus derivadas parciales quedan dadas, usando la regla del cuociente, por las siguientes expresiones

$$\frac{\partial g}{\partial x}(x,y) = \frac{y(x^2+y^2+1) - xy \cdot 2x}{(x^2+y^2+1)^2} = \frac{y^3+y-x^2y}{(x^2+y^2+1)^2}$$
$$\frac{\partial g}{\partial y}(x,y) = \frac{x(x^2+y^2+1) - xy \cdot 2y}{(x^2+y^2+1)^2} = \frac{x^3+x-xy^2}{(x^2+y^2+1)^2}$$

$\diamondsuit$

En el caso de que estudiemos una función $f : \Omega \in \mathbb{R}^n \to \mathbb{R}^m$, lo que haremos será calcular las derivadas parciales de cada una de sus funciones componentes $f_i(x)$, $i = 1, \ldots, m$. Para escribirlas de forma resumida, definimos la siguiente noción.

Definición 2.1.6 (Matriz Jacobiana)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función, con Ω un conjunto abierto y $x_0 \in \Omega$. Llamaremos Matriz Jacobiana de f en x_0, a la matriz $Df(x_0) \in \mathcal{M}_{m\times n}(\mathbb{R})$ dada por*

$$Df(x_0)_{ij} = \frac{\partial f_i}{\partial x_j}(x_0).$$

Ejemplo 2.1.7. Consideremos un ejemplo que ayude a orientarnos mejor en las dimensiones de la Matriz Jacobiana y sus respectivos elementos.

Sea la función $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$f(x,y) = \begin{pmatrix} 3xy \\ -\operatorname{sen}(x^2) \\ y \cdot \operatorname{sen}(y) \end{pmatrix}.$$

Como $f : \mathbb{R}^2 \to \mathbb{R}^3$, entonces para cualquier $(x_0, y_0) \in \mathbb{R}^2$ tenemos que $\underline{Df(x_0, y_0) \text{ es una matriz de } 3 \times 2.}$ Es decir, avazamos hacia abajo por cada *función componente* y avanzamos hacia la derecha por cada *variable.*
La Tabla 2.1 muestra la "distribución espacial" de las derivadas parciales.
Podemos entonces escribir la Matriz Jacobiana asociada a la función f en el punto (x_0, y_0) por:

$$Df(x_0,y_0) = \begin{bmatrix} 3y_0 & 3x_0 \\ -2x_0\cos(x_0^2) & 0 \\ {}^1/_{y_0} & \operatorname{sen}(y_0) + y_0\cos(y_0) \end{bmatrix}.$$

$\diamondsuit$

	x	y
f_1	$\dfrac{\partial f_1}{\partial x}(x_0, y_0) = 3y_0$	$\dfrac{\partial f_1}{\partial y}(x_0, y_0) = 3x_0$
f_2	$\dfrac{\partial f_2}{\partial x}(x_0, y_0) = -\cos(x_0^2) \cdot 2x_0$	$\dfrac{\partial f_2}{\partial y}(x_0, y_0) = 0$
f_3	$\dfrac{\partial f_3}{\partial x_0}(x_0, y_0) = 0$	$\dfrac{\partial f_3}{\partial y}(x_0, y_0) = \operatorname{sen}(y_0) + y_0 \cdot \cos(y_0)$

Cuadro 2.1: Elementos matriz jacobiana

La matriz jacobiana es la que nos permitirá más adelante definir la derivada de una función. En consecuencia, será particularmente interesante obtener propiedades sobre su comportamiento. Demostremos un primer resultado elemental.

Proposición 2.1.8 (Álgebra de Jacobianos). *Sean $f, g : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ funciones y $x_0 \in \Omega$ tales que las matrices jacobianas $Df(x_0)$ y $Dg(x_0)$ existen. Entonces,*

1. $D(f+g)(x_0) = Df(x_0) + Dg(x_0)$.
2. *Si* $\lambda \in \mathbb{R}$, *entonces* $D(\lambda f)(x_0) = \lambda Df(x_0)$.
3. *Si* $m = 1$, *entonces*
$$D(f \cdot g)(x_0) = g(x_0)Df(x_0) + f(x_0)Dg(x_0).$$
4. *Si* $m = 1$, *entonces*
$$D\left(\frac{1}{f}\right)(x_0) = -\frac{1}{f^2(x_0)}Df(x_0).$$

Demostración. Cada una de las proposiciones es una *igualdad de matrices*, por lo que se usará la misma estrategia cada vez: demostrar que las matrices son iguales coordenada a coordenada.

1. Jacobiano de la suma.
   - Sean $i \in [m]$ y $j \in [n]$ cualquiera.
   - Por definición de la matriz Jacobiana:
   $$D(f+g)(x_0)_{ij} = \frac{\partial (f+g)_i}{\partial x_j}(x_0),\ Df(x_0)_{ij} = \frac{\partial f_i}{\partial x_j}(x_0),\ Dg(x_0)_{ij} = \frac{\partial g_i}{\partial x_j}(x_0).$$
   - La derivada parcial cumple la regla de la suma (porque la hereda de las derivadas en $\mathbb{R}$), por lo que
   $$\frac{\partial (f+g)_i}{\partial x_j}(x_0) = \frac{\partial f_i}{\partial x_j}(x_0) + \frac{\partial g_i}{\partial x_j}(x_0).$$
   - Uniendo los puntos anteriores, concluimos que
   $$D(f+g)(x_0)_{ij} = Df(x_0)_{ij} + Dg(x_0)_{ij}.$$
   - Como i, j eran arbitrarios, la igualdad se cumple para toda la matriz. En consecuencia,
   $$D(f+g)(x_0) = Df(x_0) + Dg(x_0).$$

2. Jacobiano de la ponderación por escalar.

   - Sean $i \in [m]$ y $j \in [n]$ cualquiera.
   - Por definición de la matriz Jacobiana:

   $$D(\lambda f)(x_0)_{ij} = \frac{\partial \lambda f_i}{\partial x_j}(x_0), \; Df(x_0)_{ij} = \frac{\partial f_i}{\partial x_j}(x_0)$$

   - De forma similar, la derivada parcial cumple la regla de la ponderación (porque la hereda de las derivadas en $\mathbb{R}$)

   $$\frac{\partial \lambda f_i}{\partial x_j}(x_0) = \lambda \frac{\partial f_i}{\partial x_j}(x_0).$$

   - Uniendo ambos puntos obtenemos:

   $$D(\lambda f)(x_0)_{ij} = \lambda Df(x_0)_{ij}$$

   - Como i, j eran arbitrarios, la igualdad se cumple para toda la matriz. En consecuencia,

   $$D(\lambda f)(x_0) = \lambda Df(x_0).$$

3. Jacobiano del producto.

   - Sea $j \in [n]$ cualquiera (en este caso $i = 1$ siempre).
   - Por definición de la matriz Jacobiana:

   $$D(f \cdot g)(x_0)_{1j} = \frac{\partial (f \cdot g)}{\partial x_j}(x_0), \; Df(x_0)_{1j} = \frac{\partial f}{\partial x_j}(x_0), \; Dg(x_0)_{1j} = \frac{\partial g}{\partial x_j}(x_0),$$

   - De forma similar, la derivada parcial cumple la regla de producto (porque la hereda de las derivadas en $\mathbb{R}$)

   $$\frac{\partial (f \cdot g)}{\partial x_j}(x_0) = g(x_0)\frac{\partial f}{\partial x_j}(x_0) + f(x_0)\frac{\partial g}{\partial x_j}(x_0).$$

   - Uniendo los puntos anteriores se obtiene:

   $$D(f \cdot g)(x_0)_{1j} = g(x_0)Df(x_0)_{1j} + f(x_0)Dg(x_0)_{1j}.$$

   - Por igualdad de cada coordenada de la matriz, se obtiene el resultado.

4. Jacobiano de la inversa. Esta demostración se resuelve de forma equivalente a las anteriores, y queda como ejercicio para el/la lector/a.

□

2.2 La Derivada en $\mathbb{R}^n$

Como mencionamos en la sección anterior, en el caso de las funciones reales $f : \mathbb{R} \to \mathbb{R}$, la derivada cumple con ser la pendiente de la recta tangente al punto estudiado. En otras palabras, la tangente a la función f en x_0 corresponde a la recta de pendiente $m = f'(x_0)$, que además pasa por el punto $(x_0, f(x_0))$. Si calculamos la ecuación de esta recta, obtenemos

$$\frac{y - f(x_0)}{x - x_0} = f'(x_0) \quad \iff \quad y = f(x_0) + f'(x_0)(x - x_0).$$

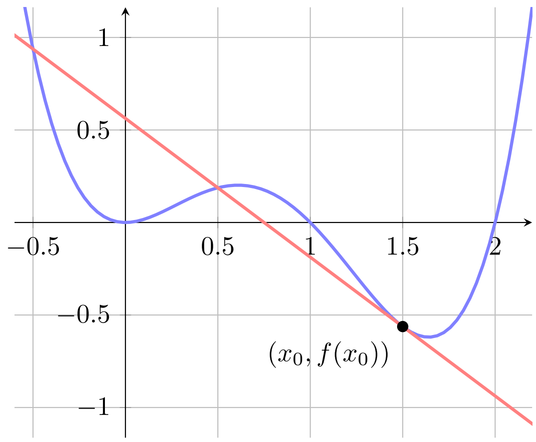


Figura 2.3: En rojo, la recta $y = f(x_0) + f'(x_0)(x - x_0)$ para $f(x) = x^2(x-1)(x-2)$ y $x_0 = \frac{3}{2}$.

En la Figura (2.3), es posible apreciar que "cerca" de x_0, la recta tangente y la función *se parecen*. Esta noción de similaridad se puede escribir de la siguiente manera: para $\delta > 0$ suficientemente pequeño y $x \in (x_0 - \delta, x_0 + \delta)$

$$f(x) \sim f(x_0) + f'(x_0)(x - x_0).$$

Para las funciones en varias variables $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $x_0 \in \Omega$, buscaremos construir la noción de derivada que nos permita obtener un resultado similar de aproximación de la función. En otras palabras, una derivada tal que

$$f(x) \sim f(x_0) + Df(x_0)(x - x_0),$$

cuando x está *cerca* de x_0.
La notación $Df(x_0)$ no es azarosa, lo que estamos buscando es que la matriz jacobiana $Df(x_0)$ cumpla el rol de la derivada en este nuevo escenario. Una pregunta natural en este punto corresponde a la siguiente: ¿será que el jacobiano, efectivamente cumple esta propiedad?

Consideremos el cambio de variables $x = x_0 + h$, y reescribamos la expresión anterior como una igualdad

$$f(x_0 + h) = f(x_0) + Df(x_0)h + o(h), \tag{2.3}$$

donde hemos definido la función $o : \mathbb{R}^n \to \mathbb{R}^m$, para que represente el error de nuestra aproximación, en función de h. En principio, nos gustaría que $o(h) \to 0$ cuando $h \to 0$, pues de esta manera, cuando estemos cerca de x_0, el error será pequeño.
Además, recordemos que en $\mathbb{R}$, cuando la derivada existe, cumple la Ecuación (2.1), que nos entrega la motivación fundamental de la siguiente definición.

Definición 2.2.1 (Derivada)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con Ω abierto y $x_0 \in \Omega$. Decimos que f es diferenciable en x_0 si todas sus derivadas parciales existen, y además cumplen*

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} = 0. \tag{2.4}$$

*Cuando esto ocurre, decimos que $Df(x_0)$ es **la derivada de f en** x_0.*

La interpretación de esta definición se fundamenta en la Ecuación (2.3), observando que cuando f es diferenciable, podemos escribir la igualdad

$$f(x_0 + h) = f(x_0) + Df(x_0)h + \underbrace{f(x_0 + h) - f(x_0) - Df(x_0)h}_{=o(h)},$$

donde hacemos la identificación $o(h) = f(x_0 + h) - f(x_0) - Df(x_0)h$, que corresponde al numerador del límite planteado en la Ecuación (2.4). Cuando pedimos que $\frac{\|o(h)\|}{\|h\|} \to 0$, estamos pidiendo tal como se discutió anteriormente, que el error tienda a cero cuando h hace lo propio, con una observación adicional: que $o(h) \to 0$ **más rápido** que h.

Ejemplo 2.2.2. Consideremos la función dada por $f(x, y, z) = x^2 + y^2 + z^2$, y probemos que es diferenciable en cualquier $(x, y, z) \in \mathbb{R}^3$.

1) Calculemos la Matriz Jacobiana de f: $Df(x, y, z) = \begin{bmatrix} 2x & 2y & 2z \end{bmatrix}$.

2) Usando $h = (h_1, h_2, h_3) \in \mathbb{R}^3$, calculemos el producto punto:

$$Df(x, y, z) \cdot (h_1, h_2, h_3) = \begin{bmatrix} 2x & 2y & 2z \end{bmatrix} \cdot (h_1, h_2, h_3) = 2xh_1 + 2yh_2 + 2zh_3.$$

3) Reemplazando en (2.4)

$$\begin{aligned} & \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} \\ = & \frac{\|f(x + h_1, y + h_2, z + h_3) - f(x, y, z) - Df(x, y, z) \cdot (h_1, h_2, h_3)\|}{\|(h_1, h_2, h_3)\|} \\ = & \frac{|(x + h_1)^2 + (y + h_2)^2 + (z + h_3)^2 - (x^2 + y^2 + z^2) - (2xh_1 + 2yh_2 + 2zh_3)|}{\sqrt{h_1^2 + h_2^2 + h_3^2}} \\ = & \frac{h_1^2 + h_2^2 + h_3^2}{\sqrt{h_1^2 + h_2^2 + h_3^2}} \\ = & \sqrt{h_1^2 + h_2^2 + h_3^2}. \end{aligned}$$

4) Finalmente, tomando límite:

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} = \lim_{(h_1, h_2, h_3) \to (0,0,0)} \sqrt{h_1^2 + h_2^2 + h_3^2} = 0.$$

5) Concluimos que f es una función diferenciable en $(x, y, z) \in \mathbb{R}^3$, y su derivada corresponde a $Df(x, y, z) = \begin{bmatrix} 2x & 2y & 2z \end{bmatrix}$.

$\Diamond$

Ejemplo 2.2.3. Consideremos ahora la función definida por ramas, dada por

$$g(x, y) = \begin{cases} \dfrac{xy \operatorname{sen}(x)}{x^2 + y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}.$$

Estudiemos la diferenciabilidad <u>en el origen</u>.

1) Calculemos la Matriz Jacobiana de g en $(0, 0)$:

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

por lo que $Dg(0, 0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$.

2) Usando $h = (h_1, h_2) \in \mathbb{R}^2$, dado lo anterior, tenemos que $Dg(0, 0) \cdot (h_1, h_2) = 0$.

3) Reemplazando en (2.4)

$$\begin{aligned}
&\frac{\|g(x_0+h)-g(x_0)-Dg(x_0)h\|}{\|h\|}\\
=\ &\frac{\|g(h_1,h_2)-g(0,0)-Dg(0,0)\cdot(h_1,h_2)\|}{\|(h_1,h_2)\|}\\
=\ &\left|\frac{h_1h_2\operatorname{sen}(h_1)}{(h_1^2+h_2^2)^{3/2}}\right|
\end{aligned}$$

4) En vez de calcular el límite con $(h_1,h_2)\to(0,0)$, usemos una sucesión en particular. Consideremos $h_{1,k}=h_{2,k}=\frac{1}{k}$, entonces tenemos:

$$\frac{\frac{1}{k}\cdot\frac{1}{k}\operatorname{sen}\left(\frac{1}{k}\right)}{\left(\frac{1}{k^2}+\frac{1}{k^2}\right)^{3/2}}=\frac{\frac{1}{k^2}\operatorname{sen}\left(\frac{1}{k}\right)}{\frac{2^{3/2}}{k^3}}=\frac{1}{2^{3/2}}\cdot\frac{\operatorname{sen}\left(\frac{1}{k}\right)}{\frac{k^2}{k^3}}=\frac{1}{2\sqrt{2}}\cdot\frac{\operatorname{sen}\left(\frac{1}{k}\right)}{\frac{1}{k}}.$$

5) Usando el límite conocido $\lim_{x\to 0}\left(\frac{\operatorname{sen} x}{x}\right)=1$, tenemos que, cuando $k\to\infty$, el resultado anterior tiende a $\frac{1}{2\sqrt{2}}$.

6) Concluimos que, de existir el límite, este no es cero (lo que necesitamos de (2.4)), por lo que g **no es diferenciable en** $(0,0)$, a pesar que su jacobiano sí existe en ese punto.

$\diamondsuit$

Un primer resultado interesante, que combina las ideas de diferenciabilidad y continuidad previamente estudiadas, corresponde al siguiente resultado.

Proposición 2.2.4. *Sea $f:\Omega\subseteq\mathbb{R}^n\to\mathbb{R}^m$ una función, con Ω un conjunto abierto y $x_0\in\Omega$. Si f es diferenciable en x_0, entonces f es continua en x_0.*

Demostración. Similar al capítulo anterior, esta demostración la haremos en "sentido inverso" (ver la demostración de la Proposición 1.2.11).

- Sea $x_0\in\Omega$ tal que f es diferenciable en x_0 (hipótesis).
- Queremos demostrar que f es continua en x_0.
- Usando la Definición 1.5.10 queremos que $\lim_{x\to x_0} f(x)=f(x_0)$.
- Haciendo el cambio de variable $x=x_0+h$, queremos que $\lim_{h\to 0} f(x_0+h)=f(x_0)$.
- Es equilavente a $\lim_{h\to 0}\|f(x_0+h)-f(x_0)\|=0$.
- Busquemos acotar $\|f(x_0+h)-f(x_0)\|$:

$$\begin{aligned}
\|f(x_0+h)-f(x_0)\|&=\|f(x_0+h)-f(x_0)-Df(x_0)h+Df(x_0)h\|\\
&\leq\|f(x_0+h)-f(x_0)-Df(x_0)h\|+\|Df(x_0)h\|\\
&\leq\|f(x_0+h)-f(x_0)-Df(x_0)h\|+\|Df(x_0)\|\cdot\|h\|,
\end{aligned}$$

  donde en la última desigualdad usamos el Lema 2.2.5.
- Tomando límite, y usando la hipótesis de diferenciabilidad, tenemos que

$$\lim_{h\to 0}\|f(x_0+h)-f(x_0)\|=0.$$

- Podemos concluir que f es continua en x_0.

$\square$

Lema 2.2.5. *Sea $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ una matriz y $x \in \mathbb{R}^n$ un vector. Considerando la norma de Frobenius, definida por:*

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

Entonces tenemos que

$$\|Mx\| \leq \|M\| \cdot \|x\|.$$

Demostración. Antes de calcular la demostración, vale la pena detenerse en los elementos de este lema y sus dimensiones.

- $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ es una matriz con m filas y n columnas. Llamaremos a cada elemento de la matriz a_{ij} con $i \in [m]$ (filas) y $j \in [n]$ (columnas).
- x es un vector en $\mathbb{R}^n$, para ser consistentes con los subíndices los elementos de x serán x_j con $j \in [n]$.
- Dada la multiplicación de matrices por vector, tenemos que Mx es un vector de $\mathbb{R}^m$ y sus elemtos serán $(Mx)_i$ con $i \in [m]$.

Por último, es importante recordar que para un vector $x \in \mathbb{R}^n :\ x_j^2 \leq \|x\|$ para todo $j \in [n]$.
Ahora sí, podemos calcular la desigualdad.

$$\begin{aligned}
\|Mx\| = \sqrt{\sum_{i=1}^{m}(Mx)_i^2} &= \sqrt{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\cdot x_j\right)^2} \\
&\leq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij}\cdot x_j)^2} \\
&= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\cdot x_j^2} \\
&\leq \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij}\cdot \|x\|)^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}\cdot\|x\| = \|M\|\cdot\|x\|
\end{aligned}$$

$\square$

2.2.1 Criterios de Diferenciabilidad

Como ya hemos visto a partir de los ejemplos desarrollados en la sección anterior, en general, demostrar diferenciabilidad de una función es difícil. Por ende, se vuelve deseable contar con algunas herramientas que permitan discernir, sin la necesidad de verificar la Ecuación (2.4), si la función que estamos estudiando se trata o no, de una función diferenciable.

Teorema 2.2.6 (Primer Criterio de Diferenciabilidad)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función y $x_0 \in \Omega$, tales que se verifica*

1. *Existe un conjunto abierto $A \subseteq \Omega$, con $x_0 \in A$, tal que para todo $x \in A$, todas las derivadas parciales de f existen.*

2. *Para todo* $i \in [m]$ *y* $j \in [n]$*, la función*

$$\frac{\partial f_i}{\partial x_j} : A \to \mathbb{R}^m$$

es continua en x_0*.*

Entonces, f *es* ***diferenciable*** *en* x_0*.*

Demostración. Esta demostración será desarrollada en detalle para el caso $n = 2$, luego el mismo esquema de demostración será utilizado para una demostración en el caso general.

Caso $n = 2$**.**

- Comencemos con $m = 1$. Sea $f : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$ y $(x_0, y_0) \in \Omega$.
- Queremos demostrar que f es diferenciable en (x_0, y_0) usando la continuidad de las derivadas parciales en (x_0, y_0).
- Sea $\varepsilon > 0$.
- Por continuidad, existen $\delta_1, \delta_2 > 0$ tales que

$$\|(x_0, y_0) - (x, y)\| \leq \delta_1 \implies \left|\frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x_0, y_0)\right| \leq \varepsilon \tag{2.5}$$

$$\|(x_0, y_0) - (x, y)\| \leq \delta_2 \implies \left|\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y_0)\right| \leq \varepsilon. \tag{2.6}$$

- Tomemos $\delta = \text{mín}\{\delta_1, \delta_2\}$ y $h = (h_1, h_2) \in \mathbb{R}^2$ tal que $\|h\| < \delta$.
- Estudiemos el númerador de la ecuación (2.4).

$$\begin{aligned}
&f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - Df(x_0, y_0) \cdot (h_1, h_2) \\
=&f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0) \cdot h_1 - \frac{\partial f}{\partial y}(x_0, y_0) \cdot h_2 \\
=&f(x_0 + h_1, y_0 + h_2) + f(x_0, y_0 + h_2) - f(x_0, y_0 + h_2) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h_1 - \frac{\partial f}{\partial y}(x_0, y_0)h_2 \\
=&\underbrace{f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0 + h_2)}_{(1)} - \frac{\partial f}{\partial x}(x_0, y_0)h_1 + \underbrace{f(x_0, y_0 + h_2) - f(x_0, y_0)}_{(2)} - \frac{\partial f}{\partial y}(x_0, y_0)h_2.
\end{aligned}$$

(1) Dejando $y_0 + h_2$ fijos, definamos $g_1 : \ [0, h_1] \to \mathbb{R}$ dada por

$$g_1(t) = f(x_0 + t, y_0 + h_2).$$

Tenemos que g_1 es continua y diferenciable, con derivada:

$$g_1'(t) = \frac{dg_1(t)}{dt} = \frac{\partial f}{\partial x}(x_0 + t, y_0 + h_2)$$

Ahora podemos aplicar el Teorema del Valor Medio (TVM), por lo que existe $c_1 \in [0, h_1]$ tal que

$$\begin{aligned}
g_1'(c_1) &= \frac{g_1(h_1) - g_1(0)}{h_1} \\
\iff \frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2) &= \frac{f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0 + h_2)}{h_1} \\
\iff \frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2) \cdot h_1 &= f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0 + h_2).
\end{aligned}$$

(2) De forma análoga, dejando x_0 fijo, definamos $g_2 : [0, h_2] \to \mathbb{R}$ dada por

$$g_2(t) = f(x_0, y_0 + t).$$

Como g_2 es continua y diferenciable podemos aplicar TVM, por lo que existe $c_2 \in [0, h_2]$ tal que

$$g_2'(c_2) = \frac{g_2(h_2) - g_2(0)}{h_2} \iff \frac{\partial f}{\partial y}(x_0, y_0 + c_2) \cdot h_2 = f(x_0, y_0 + h_2) - f(x_0, y_0).$$

- Reemplazando (1) y (2) y recordando que podemos acotar $|h_i| \leq \|h\|$, desarrollemos la ecuación (2.4)

$$\begin{aligned}
& \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} \\
= \quad & \frac{|f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - Df(x_0, y_0)(h_1, h_2)|}{\|h\|} \\
= \quad & \frac{\left|\frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2)h_1 - \frac{\partial f}{\partial x}(x_0, y_0)h_1 + \frac{\partial f}{\partial y}(x_0, y_0 + c_2)h_2 - \frac{\partial f}{\partial y}(x_0, y_0)h_2\right|}{\|h\|} \\
\leq \quad & \frac{\left|\frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2)\cancel{\|h\|} - \frac{\partial f}{\partial x}(x_0, y_0)\cancel{\|h\|} + \frac{\partial f}{\partial y}(x_0, y_0 + c_2)\cancel{\|h\|} - \frac{\partial f}{\partial y}(x_0, y_0)\cancel{\|h\|}\right|}{\cancel{\|h\|}} \\
= \quad & \left|\frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2) - \frac{\partial f}{\partial x}(x_0, y_0) + \frac{\partial f}{\partial y}(x_0, y_0 + c_2) - \frac{\partial f}{\partial y}(x_0, y_0)\right| \\
\leq \quad & \left|\frac{\partial f}{\partial x}(x_0 + c_1, y_0 + h_2) - \frac{\partial f}{\partial x}(x_0, y_0)\right| + \left|\frac{\partial f}{\partial y}(x_0, y_0 + c_2) - \frac{\partial f}{\partial y}(x_0, y_0)\right| \\
\leq \quad & \varepsilon + \varepsilon \\
= \quad & 2\varepsilon.
\end{aligned}$$

La última desigualdad se obtiene por la continuidad de las derivadas parciales, como establecido en (2.5) y (2.6). Es importante, y queda propuesto al lector, verificar que se cumplen las hipótesis de dichas ecuaciones.

- Finalmente, como ε es arbitrario, podemos concluir que

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} = 0,$$

por lo que f es diferenciable en (x_0, y_0).

- Para $m > 1$, se puede aplicar el resultado a cada una de sus funciones coordenadas. En virtud del Teorema 1.5.5, esto es suficiente para probar la existencia y valor nulo de este límite.

Caso general.

- Comencemos con $m = 1$. Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ y $x_0 \in \Omega$.
- Queremos demostrar que f es diferenciable en x_0 <u>usando la continuidad de las derivadas parciales</u> en x_0.
- Sea $\varepsilon > 0$.

- Por continuidad, podemos elegir $\delta > 0$ suficientemente pequeño tal que $\forall h \in \mathbb{R}^n$ con $\|h\| < \delta$:

$$\forall i \in [n]: \quad \left| \frac{\partial f}{\partial x_i}(x_0 + h) - \frac{\partial f}{\partial x_i}(x_0) \right| < \varepsilon. \tag{2.7}$$

- Tomemos $h \in \mathbb{R}^n$ tal que $\|h\| < \delta$. Para facilitar la notación definamos para $i \in [n]$ el vector $\phi_i(h) = (h_1, \ldots, h_i, 0, \ldots, 0)$, donde además será $\phi_0(h) = 0 \in \mathbb{R}^n$.

- Usando suma telescópica podemos escribir:

$$\begin{aligned} & f(x_0 + h) - f(x_0) \\ = \; & f(x_0 + \phi_n(h)) - f(x_0 + \phi_0(h)) \\ = \; & \sum_{i=1}^{n} f(x_0 + \phi_i(h)) - f(x_0 + \phi_{i-1}(h)) \end{aligned}$$

- Se puede aplicar TVM a cada elemento de la suma, pues representa "movimiento" en una sola dimensión. Entonces para cada $i \in [n]$, $\exists c_i \in [x_0 + \phi_{i-1}(h), x_0 + \phi_i(h)]$ tal que

$$f(x_0 + \phi_i(h)) - f(x_0 + \phi_{i-1}(h)) = \frac{\partial f}{\partial x_i}(c_i) \cdot (\phi_i(h) - \phi_{i-1}(h)) = \frac{\partial f}{\partial x_i}(c_i) h_i.$$

- Ahora, buscando acotar y usando los puntos anteriores, desarrollemos la ecuación (2.4).

$$\begin{aligned} & \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} \\ = \; & \frac{\left| f(x_0 + h) - f(x_0) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) h_i \right|}{\|h\|} \\ = \; & \frac{\left| \sum_{i=1}^{n} f(x_0 + \phi_i(h)) - f(x_0 + \phi_{i-1}(h)) - \frac{\partial f}{\partial x_i}(x_0) h_i \right|}{\|h\|} \\ = \; & \frac{\left| \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(c_i) h_i - \frac{\partial f}{\partial x_i}(x_0) h_i \right|}{\|h\|} \\ \leq \; & \frac{\sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(c_i) h_i - \frac{\partial f}{\partial x_i}(x_0) h_i \right|}{\|h\|} \\ \leq \; & \frac{\sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(c_i) - \frac{\partial f}{\partial x_i}(x_0) \right| \cdot \cancel{\|h\|}}{\cancel{\|h\|}} \\ = \; & \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(c_i) - \frac{\partial f}{\partial x_i}(x_0) \right| \\ \leq \; & \sum_{i=1}^{n} \varepsilon \\ = \; & n \cdot \varepsilon. \end{aligned}$$

  La última desigualdad se obtiene por la continuidad de las derivadas parciales, como establecido en (2.7) (verificando que $\|c_i - x_0\| \leq \|(x_0 + h) - x_0\| = \|h\| < \delta$ para cada $i \in [n]$).

- Finalmente, como ε es arbitrario, podemos concluir que

$$\lim_{h \to 0} \frac{\|f(x_0 + h) - f(x_0) - Df(x_0)h\|}{\|h\|} = 0,$$

  por lo que f es diferenciable en (x_0, y_0).

- Para $m > 1$, se puede aplicar el resultado a cada una de sus funciones coordenadas. En virtud del Teorema 1.5.5, esto es suficiente para probar la existencia y valor nulo de este límite.

$\square$

Ejemplo 2.2.7. Usando el teorema anterior, podemos mostrar directamente que la función

$$f(x,y) = \begin{pmatrix} xy \\ x^2y \end{pmatrix}$$

es diferenciable en todo $\mathbb{R}^2$. Para ello, basta con notar que

$$Df(x,y) = \begin{bmatrix} y & x \\ 2xy & x^2 \end{bmatrix}$$

y dado que todas sus componentes son continuas en $\mathbb{R}^2$, deducimos el resultado. $\Diamond$

Atención

El teorema anterior sirve cuando las derivadas parciales existen en todo punto. Por ende, surge una pregunta natural: ¿qué pasa cuando la matriz jacobiana no existe en algún punto?

Estudiemos esta pregunta, analizando la función

$$g(x,y) = \begin{cases} x^2+y^2 & \text{si } x^2+y^2 \le 1 \\ 1 & \text{si } x^2+y^2 > 1 \end{cases},$$

cuyo gráfico está dado por lo mostrado en la Figura (2.4).

Consideremos los conjuntos

- $A_1 = B(0,1)$,
- $A_2 = \mathbb{R}^2 \setminus \overline{B}(0,1)$, y
- $A_3 = \overline{B}(0,1) \setminus B(0,1)$.

Notemos que en A_1, las derivadas parciales son continuas:

$$\frac{\partial g}{\partial x}(x,y) = 2x \qquad \text{y} \qquad \frac{\partial g}{\partial y}(x,y) = 2y.$$

Por lo tanto g es diferenciable en todo punto de A_1 aunque no sea diferenciable en la frontera (que corresponde al conjunto A_3).

Por lo demás, las derivadas parciales en A_2 también son continuas –al ser constantes igual a cero– por lo que g también es diferenciable en A_2.

Por otra parte, aprovechando lo que ya sabemos sobre álgebra de matrices jacobianas, podemos obtener un segundo criterio para estudiar diferenciabilidad: cuando sepamos de antemano que una o más funciones son diferenciables, entonces las operaciones básicas entre ellas también lo serán.

Teorema 2.2.8 (Segundo Criterio de Diferenciabilidad)**.** *Sean $f, g : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ dos funciones y $x_0 \in \Omega$ tal que f y g son diferenciables en x_0. Entonces:*

1. *$f + g$ es diferenciable en x_0.*

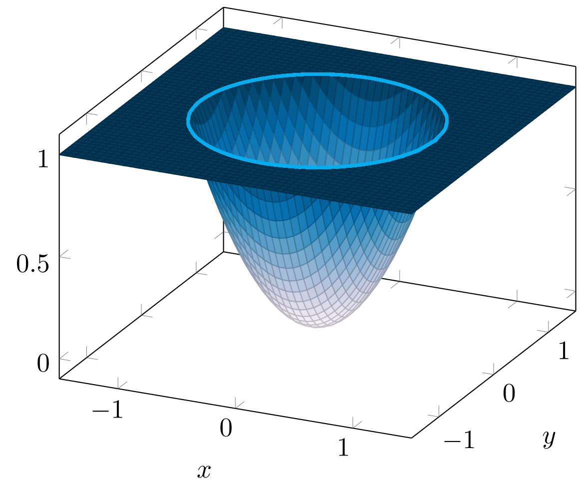


Figura 2.4: Gráfico de $g(x, y)$.

2. *Para todo $\lambda \in \mathbb{R}$, (λf) es diferenciable en x_0.*
3. *Si $m = 1$, $f \cdot g$ es diferenciable en x_0.*
4. *Si $m = 1$ y* ***$1/f$ está bien definida en un abierto*** *A, $1/f$ es diferenciable en x_0.*

Demostración. Demostramos cada apartado por separado. La idea cada vez es la misma:

- Escribir la Ecuación de la Derivada (2.4).
- Utilizar álgebra de Jacobianos (Teorema 2.1.8) para que aparezcan los Jacobianos de f y/o g.
- "Armar" la Ecuación (2.4) para f y/o g.
- Utilizar que f y g son diferenciables para que ese término tienda a cero.
- Analizar el resto de términos que queden.

1. Diferenciabilidad de la suma. Usando el esquema de demostración anterior tenemos que:

$$\begin{aligned}
& \frac{\|(f+g)(x_0+h)-(f+g)(x_0)-D(f+g)(x_0)h\|}{\|h\|} \\
= \quad & \frac{\|(f+g)(x_0+h)-(f+g)(x_0)-Df(x_0)h-Dg(x_0)h\|}{\|h\|} \\
= \quad & \frac{\|f(x_0+h)+g(x_0+h)-f(x_0)-g(x_0)-Df(x_0)h-Dg(x_0)h\|}{\|h\|} \\
& \frac{\|\,(f(x_0+h)-f(x_0)-Df(x_0)h)+(g(x_0+h)-g(x_0)-Dg(x_0)h)\,\|}{\|h\|} \\
\leq \quad & \underbrace{\frac{\|f(x_0+h)-f(x_0)-Df(x_0)h\|}{\|h\|}}_{\to 0}+\underbrace{\frac{\|g(x_0+h)-g(x_0)-Dg(x_0)h\|}{\|h\|}}_{\to 0} \\
\longrightarrow \quad & 0.
\end{aligned}$$

Como se cumple que el límite es cero en la Ecuación (2.4), queda demostrado que $f+g$ es diferenciable en x_0.

2. Diferenciabilidad de la ponderación. Usando el esquema de demostración anterior tenemos que:

$$\begin{aligned}
& \frac{\|(\lambda f)(x_0+h)-(\lambda f)(x_0)-D(\lambda f)(x_0)h\|}{\|h\|} \\
= \quad & \frac{\|(\lambda f)(x_0+h)-(\lambda f)(x_0)-\lambda \cdot Df(x_0)h\|}{\|h\|} \\
= \quad & \frac{\|\lambda \cdot f(x_0+h)-\lambda \cdot f(x_0)-\lambda \cdot Df(x_0)h\|}{\|h\|} \\
= \quad & |\lambda| \cdot \underbrace{\frac{\|f(x_0+h)-f(x_0)-Df(x_0)h\|}{\|h\|}}_{\to 0} \\
\longrightarrow \quad & 0.
\end{aligned}$$

Como se cumple que el límite es cero en la Ecuación (2.4), queda demostrado que λf es diferenciable en x_0.

3. Diferenciabilidad de la multiplicación ($m=1$). Usaremos el mismo esquema de demostración, pero al ser más largo iremos analizando la ecuación por partes.

   - Comenzamos igual:

$$\begin{aligned}
& \frac{\|(f\cdot g)(x_0+h)-(f\cdot g)(x_0)-D(f\cdot g)(x_0)h\|}{\|h\|} \\
= \quad & \frac{|(f\cdot g)(x_0+h)-(f\cdot g)(x_0)-g(x_0)Df(x_0)h-f(x_0)Dg(x_0)h|}{\|h\|} \\
= \quad & \frac{|f(x_0+h)g(x_0+h)-f(x_0)g(x_0)-g(x_0)\mathbf{Df(x_0)h}-f(x_0)Dg(x_0)h|}{\|h\|}
\end{aligned}$$

   - Identificando el Jacobiano de f, vemos primero que está acompañado de $g(x_0)$. También vemos que $g(x_0)$ acompaña a $f(x_0)$, entonces juntamos los elementos que nos permiten construir la ecuación para f y agregamos los necesarios (“ni quita ni pone”).

$$\begin{aligned}
& \frac{|f(x_0+h)g(x_0+h)-f(x_0)g(x_0)-g(x_0)Df(x_0)h-f(x_0)Dg(x_0)h|}{\|h\|} \\
= \quad & \frac{|f(x_0+h)g(x_0+h)-f(x_0)g(x_0)-g(x_0)Df(x_0)h-f(x_0)Dg(x_0)h+\mathbf{f(x_0+h)g(x_0)}-\mathbf{f(x_0+h)g(x_0)}|}{\|h\|} \\
= \quad & \frac{|f(x_0+h)g(x_0)-f(x_0)g(x_0)-g(x_0)Df(x_0)h+f(x_0+h)g(x_0+h)-f(x_0+h)g(x_0)-f(x_0)Dg(x_0)h|}{\|h\|} \\
\leq \quad & \underbrace{|g(x_0)|\frac{|f(x_0+h)-f(x_0)-Df(x_0)h|}{\|h\|}}_{(1)\to 0}+\underbrace{\frac{|f(x_0+h)g(x_0+h)-f(x_0+h)g(x_0)-f(x_0)Dg(x_0)h|}{\|h\|}}_{(2)}
\end{aligned}$$

   - En (1) logramos construir la Ecuación (2.4) para f, que sabemos tiende a cero por ser f diferenciable. El límite sigue siendo cero al multiplicar por $|g(x_0)|$ al ser un valor fijo.
   - Ahora necesitamos ver que (2) también tiende a cero. Para esto tratamos de construir la Ecuación (2.4) para g y nos vamos a centrar en los elementos que tienen $f(x_0+h)$ como término común.[1]

[1] Notemos que para el caso de f construimos la Ecuación (2.4) a partir del Jacobiano de f, mientras que para g no. Esto es porque para usar $f(x_0)Dg(x_0)h$ tendríamos que agregar dos términos por “ni quita ni pone”, mientras que hacer aparecer $f(x_0+h)Dg(x_0)h$ requiere sólo un paso.

$$
\begin{aligned}
(2) = \quad & \frac{|f(x_0+h)g(x_0+h) - f(x_0+h)g(x_0) - f(x_0)Dg(x_0)h|}{\|h\|} \\
= \quad & \frac{|f(x_0+h)g(x_0+h) - f(x_0+h)g(x_0) - \mathbf{f(x_0+h)Dg(x_0)h} + \mathbf{f(x_0+h)Dg(x_0)h} - f(x_0)Dg(x_0)h|}{\|h\|} \\
\leq \quad & \underbrace{|f(x_0+h)|\frac{|g(x_0+h) - g(x_0) - Dg(x_0)h)|}{\|h\|}}_{(3)\to 0} + \underbrace{|f(x_0+h) - f(x_0)|\frac{|Dg(x_0)h|}{\|h\|}}_{(4)}
\end{aligned}
$$

- En (3) tenemos $|f(x_0+h)|$ que converge a $|f(x_0)|$ por continuidad de f. También tenemos la Ecuación (2.4) para g, que sabemos tiende a cero por ser g diferenciable. Por lo tanto, el límite cuando $h \to 0$ es $|f(x_0)| \cdot 0 = 0$.
- Finalmente necesitamos que (4) también converja a cero, para esto vamos a usar el Lema 2.2.5 para $|Dg(x_0)h| \leq \|Dg(x_0)\| \cdot \|h\|$.

$$
\begin{aligned}
(4) = \quad & |f(x_0+h) - f(x_0)|\frac{|Dg(x_0)h|}{\|h\|} \\
\leq \quad & |f(x_0+h) - f(x_0)|\frac{\|Dg(x_0)\| \cdot \cancel{\|h\|}}{\cancel{\|h\|}} \\
= \quad & \underbrace{|f(x_0+h) - f(x_0)|}_{\to 0}\|Dg(x_0)\|
\end{aligned}
$$

- El límite de (4) tiende a cero por continuidad de f, que se mantiene al ser multiplicado por un valor constante. Con esto podemos concluir que el límite es cero en la Ecuación (2.4) y queda demostrado que $f \cdot g$ es diferenciable en x_0.

4. Diferenciabilidad de la inversa ($m = 1$).

- Para facilitar el cálculo, estudiamos primero el númerador de la Ecuación (2.4):

$$
\begin{aligned}
& (1/f)(x_0+h) - (1/f)(x_0) - D(1/f)(x_0)h \\
= \quad & (1/f)(x_0+h) - (1/f)(x_0) + Df(x_0)/f^2(x_0)h \\
= \quad & \frac{1}{f(x_0+h)} - \frac{1}{f(x_0)} + \frac{Df(x_0)h}{f^2(x_0)} \\
= \quad & \frac{f(x_0) - f(x_0+h)}{f(x_0)f(x_0+h)} + \frac{Df(x_0)h}{f^2(x_0)} \\
= \quad & \frac{f(x_0) - f(x_0+h) + \mathbf{Df(x_0)h} - \mathbf{Df(x_0)h}}{f(x_0)f(x_0+h)} + \frac{Df(x_0)h}{f^2(x_0)} \\
= \quad & \frac{f(x_0) - f(x_0+h) + Df(x_0)h}{f(x_0)f(x_0+h)} - \frac{Df(x_0)h}{f(x_0)f(x_0+h)} + \frac{Df(x_0)h}{f^2(x_0)} \\
= \quad & -\frac{f(x_0+h) - f(x_0) - Df(x_0)h}{f(x_0)f(x_0+h)} + Df(x_0)h \cdot \left(\frac{1}{f^2(x_0)} - \frac{1}{f(x_0)f(x_0+h)}\right)
\end{aligned}
$$

- Reemplazando $(*)$ en la Ecuación (2.4) completa, obtenemos:

$$\begin{aligned}
&\frac{\|(1/f)(x_0+h)-(1/f)(x_0)-D(1/f)(x_0)h\|}{\|h\|}\\
=\ &\frac{\left|-\frac{f(x_0+h)-f(x_0)-Df(x_0)h}{f(x_0)f(x_0+h)}+Df(x_0)h\ \left(\frac{1}{f^2(x_0)}-\frac{1}{f(x_0)f(x_0+h)}\right)\right|}{\|h\|}\\
\leq\ &\underbrace{\frac{1}{|f(x_0)f(x_0+h)|}\cdot\frac{|f(x_0+h)-f(x_0)-Df(x_0)h|}{\|h\|}}_{(1)}+\underbrace{\frac{|Df(x_0)h|}{\|h\|}\cdot\left|\left(\frac{1}{f^2(x_0)}-\frac{1}{f(x_0)f(x_0+h)}\right)\right|}_{(2)}
\end{aligned}$$

- Para (1) tenemos que $\frac{1}{|f(x_0)f(x_0+h)|}\to\frac{1}{|f^2(x_0)|}$ porque $\frac{1}{f(x)}$ está bien definida en un abierto A con $x_0\in A$. También tenemos la Ecuación (2.4) para f, que sabemos tiende a cero porque f es diferenciable. Por lo tanto, el límite cuando $h\to 0$ es $\frac{1}{|f^2(x_0)|}\cdot 0=0$.
- Ahora necesitamos ver si (2) también converge a cero. Para esto utilizaremos el Lema 2.2.5 para $|Df(x_0)h|\leq\|Df(x_0)\|\cdot\|h\|$.

$$\begin{aligned}
(2)=\ &\frac{|Df(x_0)h|}{\|h\|}\cdot\left|\left(\frac{1}{f^2(x_0)}-\frac{1}{f(x_0)f(x_0+h)}\right)\right|\\
\leq\ &\frac{\|Df(x_0)\|\cancel{\|h\|}}{\cancel{\|h\|}}\cdot\left|\left(\frac{1}{f^2(x_0)}-\frac{1}{f(x_0)f(x_0+h)}\right)\right|\\
=\ &\|Df(x_0)\|\cdot\underbrace{\left|\left(\frac{1}{f^2(x_0)}-\frac{1}{f(x_0)f(x_0+h)}\right)\right|}_{\to 0}
\end{aligned}$$

- El límite de (2) tiende a cero por continuidad de f, que se mantiene al ser multiplicado por un valor constante. Con esto podemos concluir que el límite es cero en la Ecuación (2.4) y queda demostrado que $\frac{1}{f}$ es diferenciable en x_0.

□

Finalmente, un tercer criterio de diferenciabilidad, heredado de la caracterización de límites de varias variables demostrada en el Teorema (1.5.5), corresponde al siguiente.

Teorema 2.2.9 (Tercer Criterio de Diferenciabilidad)**.** *Sea $f:\Omega\subseteq\mathbb{R}^n\to\mathbb{R}^m$ y $x_0\in\Omega$. Entonces,*

$$f\ \textit{es diferenciable en}\ x_0\iff\forall i\in\{1,\dots,m\}\ f_i\ \textit{es diferenciable en}\ x_0,$$

donde $f_i:\Omega\to\mathbb{R}$ es la $i-$ésima función componente de f.

Atención

Para demostrar que una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ no es diferenciable, por el momento sólo tenemos las siguientes opciones.

- Probar que $Df(x_0)$ no existe.
- Encontrar una sucesión $h_k \to 0$ tal que

$$\lim_{k\to\infty} \frac{\|o(h_k)\|}{\|h_k\|} \neq 0.$$

Los tres criterios de diferenciabilidad enunciados hasta aquí, **no permiten demostrar que una función no sea diferenciable**. Sólo permiten concluir que sí lo son, tal como sus nombres lo indican. Por ejemplo, la función

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

es diferenciable en $(0,0)$, a pesar de que sus derivadas parciales no son continuas en ese punto.

A pesar de lo mencionado en el recuadro anterior respecto de los criterios estudiados hasta aquí, es posible contar con un criterio de "no-diferenciabilidad", que surge a partir del siguiente resultado.

Teorema 2.2.10. *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $x_0 \in \Omega$. Si f es diferenciable en x_0, entonces sus derivadas direccionales cumplen que*

$$\forall v \in \mathbb{R}^n, \quad f'(x_0; v) = Df(x_0)v.$$

Demostración. Como sabemos que f es diferenciable en x_0, tenemos

$$\lim_{h\to 0} \frac{\|f(x_0+h) - f(x_0) - Df(x_0)h\|}{\|h\|} = 0.$$

Por ende, tomando una sucesión cualquiera $h_k \to 0$, el límite debe ser el mismo. Elegimos para $v \in \mathbb{R}^n \setminus \{0\}$ la sucesión dada por

$$h_k = vt_k,$$

donde $(t_k)_{k\in\mathbb{N}} \subseteq R$ es una sucesión cualquiera, con $t_k \to 0$ cuando $k \to \infty$. Reemplazando, obtenemos

$$\begin{aligned}
& \lim_{k\to\infty} \frac{\|f(x_0+vt_k) - f(x_0) - Df(x_o)vt_k\|}{\|vt_k\|} = 0 \\
\iff \quad & \lim_{k\to\infty} \frac{\|f(x_0+vt_k) - f(x_0) - t_kDf(x_0)v\|}{\|v\||t_k|} = 0 \\
\iff \quad & \lim_{k\to\infty} \frac{\|f(x_0+vt_k) - f(x_0) - t_kDf(x_0)v\|}{|t_k|} = 0 \\
\iff \quad & \lim_{k\to\infty} \left\| \frac{f(x_0+vt_k) - f(x_0)}{t_k} - Df(x_0)v \right\| = 0 \\
\implies \quad & \lim_{k\to\infty} \frac{f(x_0+vt_k) - f(x_0)}{t_k} = Df(x_0)v.
\end{aligned}$$

Como t_k es una sucesión arbitraria, concluimos que $f'(x_0; v) = Df(x_0)v$, usando la Definición (2.1.2). $\square$

La contrarrecíproca del teorema anterior, nos permite contar con el siguiente resultado, que vale como criterio para determinar si una función es o no diferenciable.

Corolario 2.2.11 (Criterio de no-Diferenciabilidad)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $x_0 \in \Omega$. Si existe $v \in \mathbb{R}^n \setminus \{0\}$ tal que*

$$f'(x_0, v) \text{ no exista, o bien } f'(x_0; v) \neq Df(x_0)v,$$

*entonces f **no es** diferenciable en x_0.*

2.3 Vector Gradiente y Plano Tangente

Consideremos el caso en que estudiamos $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$, con Ω un conjunto abierto. Cuando f es diferenciable en $x_0 \in \Omega$, su derivada corresponde a una matriz de $1 \times n$, que se puede identificar con un vector. De esto, surge la siguiente definición.

Definición 2.3.1 (Vector Gradiente)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en $x_0 \in \Omega$. Definimos su* gradiente *como el vector dado por*

$$\nabla f(x_0) = Df(x_0)^\top,$$

vale decir, como aquel obtenido al trasponer el jacobiano de la función.

El vector gradiente cumple una importante propiedad, que demostramos a continuación.

Proposición 2.3.2. *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en $x_0 \in \Omega$. Entonces, $\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$ corresponde al vector unitario que maximiza la derivada direccional de f en x_0, es decir,*

$$\max_{\|v\|=1} f'(x_0; v) = f'\left(x_0, \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}\right)$$

*A esto llamamos que $\nabla f(x_0)$ es la **dirección de máximo crecimiento** de f en x_0.*

Demostración. Lo primero es entender que $f'(x_0, v)$ representa el crecimiento o decrecimiente (según su signo) de f en x_0 según la dirección v. Es por linealidad que solo se consideran los vectores v unitarios, es decir, tal que $\|v\| = 1$.

Sea $v \in \mathbb{R}^n$ tal que $\|v\| = 1$. En virtud del Teorema 2.2.10, sabemos que

$$f'(x_0; v) = Df(x_0)v = \langle \nabla f(x_0), v \rangle,$$

y por Cauchy-Schwarz,

$$f'(x_0; v) \leq |f'(x_0; v)| \leq \|\nabla f(x_0)\| \|v\| = \|\nabla f(x_0)\| = M.$$

En este punto hemos demostrado que $f'(x_0, v)$ está acotado por M para cualquier vector unitario v. El objetivo ahora es demostrar que esa cota se alcanza en la dirección de $\nabla f(x_0)$, para esto vamos a elegir el vector unitario:

$$e = \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|},$$

para notar que

$$f'(x_0; e) = \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} \right\rangle = \frac{1}{\|\nabla f(x_0)\|} \langle \nabla f(x_0), \nabla f(x_0) \rangle = \frac{\|\nabla f(x_0)\|^2}{\|\nabla f(x_0)\|} = \|\nabla f(x_0)\| = M.$$

En resumen, hemos probado que para todo vector unitario,

$$f'(x_0; v) \leq f'(x_0; e),$$

expresión que demuestra lo buscado. □

En la Figura 2.5 se representa la proposición anterior: de todos los vectores del cíulo unitario centrado en x_0 (círuclo cyan), $\frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$ es la dirección que maximiza la magnitud de los vectores del círculo magenta, que son las derivadas direccionales para esos vectores. Más de este mismo ejemplo se puede explorar desde este applet de Geogebra.

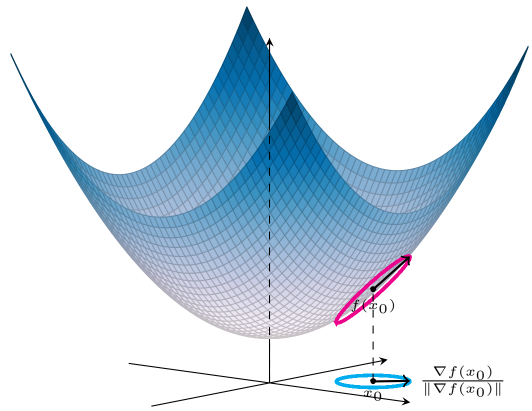


Figura 2.5: Interpretación gráfica de $\nabla f(x)$ y la dirección de máximo crecimiento para f.

Dado que estamos dando sentido geométrico a las nociones que estamos estudiando, miremos una nueva a continuación. Si tomamos $f : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$, recordemos que cerca de un punto $(x_0, y_0) \in \Omega$, es posible decir que

$$f(x, y) \sim f(x_0, y_0) + \langle \nabla f(x_0, y_0), (x - x_0, y - y_0)\rangle.$$

Es razonable pensar entonces, que al graficar la función f y de su aproximación, los grafos de ambas expresiones sean *similares*. Eso es lo que motiva la siguiente definición.

Definición 2.3.3 (Plano Tangente)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ y $x_0 \in \Omega$ tal que f es diferenciable en este punto. Definimos su* plano tangente *$H \subseteq \mathbb{R}^{n+1}$ dado por la expresión*

$$x_{n+1} = f(x_0) + \langle \nabla f(x_0), x - x_0\rangle.$$

2.4 Regla de la Cadena y Teorema del Valor Medio

En este punto de los contenidos, demostramos un resultado pendiente: la diferenciabilidad de la composición de funciones, conocida como la *regla de la cadena.*

Teorema 2.4.1 (Regla de la Cadena)**.** *Sean $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$ y $g : \Lambda \subseteq \mathbb{R}^m \to \mathbb{R}^p$ dos funciones, con Ω y Λ abiertos, y sea $x_0 \in \Omega$ tales que*

- *$f(\Omega) \subseteq \Lambda$.*
- *f es diferenciable en x_0.*
- *g es diferenciable en $f(x_0)$.*

Entonces, la composición $g \circ f : \Omega \to \mathbb{R}^p$ es diferenciable, y más aún,

$$D(g \circ f)(x_0) = Dg(f(x_0)) \cdot Df(x_0).$$

Demostración. Para demostrar lo pedido, debemos verificar que se cumple

$$\lim_{h\to 0} \frac{\|(g \circ f)(x_0 + h) - (g \circ f)(x_0) - Dg(f(x_0)) \cdot Df(x_0)h\|}{\|h\|} = 0.$$

Para facilitar la escritura, denotemos

$$k(h) = f(x_0 + h) - f(x_0), \quad y_0 = f(x_0), \quad o(h) = (g \circ f)(x_0 + h) - (g \circ f)(x_0) - Dg(f(x_0)) \cdot Df(x_0)h.$$

Como f es diferenciable, en particular es continua, por lo que $k(h) \to 0$ cuando $h \to 0$. Reajustando términos en el numerador, observamos que

$$\begin{aligned}
\frac{\|o(h)\|}{\|h\|} &= \frac{\|g(f(x_0+h)) - g(f(x_0)) - Dg(f(x_0)) \cdot Df(x_0)h\|}{\|h\|} \\
&= \frac{\|g(f(x_0+h) - f(x_0) + f(x_0)) - g(f(x_0)) - Dg(f(x_0)) \cdot (Df(x_0)h - k(h) + k(h))\|}{\|h\|} \\
&= \frac{\|g(y_0 + k(h)) - g(y_0) - Dg(y_0)k(h) + Dg(y_0)(k(h) - Df(x_0)h)\|}{\|h\|} \\
&\leq \frac{\|g(y_0 + k(h)) - g(y_0) - Dg(y_0)k(h)\| + \|Dg(y_0)\| \|f(x_0+h) - f(x_0) - Df(x_0)h\|}{\|h\|} \\
&= \underbrace{\frac{\|g(y_0 + k(h)) - g(y_0) - Dg(y_0)k(h)\|}{\|h\|}}_{(1)} + \underbrace{\frac{\|Dg(y_0)\| \|f(x_0+h) - f(x_0) - Df(x_0)h\|}{\|h\|}}_{(2)}
\end{aligned}$$

Por un lado, tenemos que (2) converge a cero por diferenciabilidad de f. Por lo tanto, solo debemos estudiar (1). Usaremos el Teorema de límite por partes (ver Teorema 1.5.9), tomando $A_1 := \{h : k(h) = 0\}$ y $A_2 = \{h : k(h) \neq 0\}$. Sea entonces $(h_j)_{j \in \mathbb{N}}$ una sucesión con $h_j \to 0$:

- **Caso 1:** Si $(h_j) \subset A_1$, entonces $(1) = 0$, y por lo tanto converge a 0 de forma trivial.
- **Caso 2:** Si $(h_j) \subset A_2$, entonces $k(h_j) \neq 0$. Más aún,

  $$\|k(h_j)\| = \|f(x_0 + h_j) - f(x_0)\| \leq \|f(x_0 + h_j) - f(x_0) - Df(x_0)h_j\| + \|Df(x_0)\| \|h_j\|.$$

  Usando esto, podemos escribir

  $$\begin{aligned}
  (1) &= \frac{\|g(y_0 + k(h_j)) - g(y_0) - Dg(y_0)k(h_j)\|}{\|h_j\|} \\
  &= \frac{\|g(y_0 + k(h_j)) - g(y_0) - Dg(y_0)k(h_j)\|}{\|k(h_j)\|} \frac{\|k(h_j)\|}{\|h_j\|} \\
  &\leq \frac{\|g(y_0 + k(h_j)) - g(y_0) - Dg(y_0)k(h_j)\|}{\|k(h_j)\|} \left(\frac{\|f(x_0 + h_j) - f(x_0) - Df(x_0)h_j\| + \|Df(x_0)\| \|h_j\|}{\|h_j\|} \right) \\
  &= \underbrace{\frac{\|g(y_0 + k(h_j)) - g(y_0) - Dg(y_0)k(h_j)\|}{\|k(h_j)\|}}_{\to 0} \underbrace{\left(\frac{\|f(x_0 + h_j) - f(x_0) - Df(x_0)h_j\|}{\|h_j\|} + \|Df(x_0)\| \right)}_{\to 0 + \|Df(x_0)\|} \xrightarrow{j \to \infty} 0.
  \end{aligned}$$

  Donde en el primer límite se usó que g es diferenciable en y_0 y que $k(h_j) \to 0$, mientras que en el segundo límite se usó la diferenciabilidad de f.

Independiente del caso, tenemos que $(1) \to 0$ cuando $j \to \infty$. Por lo tanto, concluimos que $(1) \to 0$ como función cuando $h \to 0$. Se tiene entonces que

$$\frac{\|o(h)\|}{\|h\|} \leq (1) + (2) \to 0,$$

lo que concluye la demostración. □

El teorema recién demostrado, permite obtener un corolario respecto a las derivadas parciales de una composición de funciones, analizando las componentes de la multiplicación de jacobianos del resultado.

Corolario 2.4.2 (Regla de la Cadena para Derivadas Parciales)**.** *Sean f, g como las del Teorema (2.4.1), donde $f(x) = (f_1(x), \dots, f_m(x))$. Entonces, para*

$$h = g(f_1(x), \dots, f_m(x)),$$

tenemos

$$\frac{\partial h}{\partial x_j}(x_0) = \sum_{i=1}^{m} \frac{\partial g}{\partial y_i}(f(x_0)) \frac{\partial f_i}{\partial x_j}(x_0).$$

Terminamos esta sección, demostrando un resultado teórico muy relevante para usar de forma auxiliar, desarrollando otros teoremas más adelante.

Teorema 2.4.3 (Teorema del Valor Medio)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable sobre todo Ω. Sean $x, y \in \Omega$ tales que el* segmento *entre ellos está contenido en Ω, vale decir*

$$[x, y] = \{z \in \mathbb{R}^n \ : \ z = x + \lambda(y - x), \ \lambda \in [0, 1]\} \subset \Omega.$$

Entonces, existe $c \in (0, 1)$ tal que

$$f(y) - f(x) = \langle \nabla f(x + c(y - x), y - x \rangle.$$

Demostración. Consideremos la función $\varphi : [0, 1] \to \mathbb{R}$ dada por

$$\varphi(t) = f(x + t(y - x)),$$

que cumple las hipótesis para poder aplicar el TVM en $\mathbb{R}$. En consecuencia, existe $c \in (0, 1)$ tal que

$$\varphi(1) - \varphi(0) = \varphi'(c) = Df(x + c(y - x))(y - x),$$

donde el cálculo de $\varphi'(c)$ se obtiene por regla de la cadena. Finalmente, notando que $\varphi(1) = f(y)$ y $\varphi(0) = f(x)$, se concluye lo buscado. $\square$

2.5 Derivadas de Orden Superior

Consideremos una función $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ tal que todas sus derivadas parciales existen y son continuas en Ω. Entonces, para todo $i \in \{1, \dots, m\}$ y todo $j \in \{1, \dots, n\}$ la función

$$\frac{\partial f_i}{\partial x_j} : \Omega \subset \mathbb{R}^n \to \mathbb{R}$$

podría ser también diferenciable, o al menos tener derivadas parciales. Este simple hecho, motiva la siguiente definición.

Definición 2.5.1 (Derivadas Parciales de Segundo Orden)**.** *Cuando la función $\frac{\partial f_i}{\partial x_j}$ tiene derivada parcial con respecto a x_k en un punto x_0, la anotamos como*

$$\frac{\partial^2 f_i}{\partial x_k \partial x_j}(x_0) = \frac{\partial}{\partial x_k}\left(\frac{\partial f_i}{\partial x_j}\right)(x_0).$$

En el caso particular que $k = j$, anotamos

$$\frac{\partial^2 f_i}{\partial x_j^2}(x_0) = \frac{\partial}{\partial x_j}\left(\frac{\partial f_i}{\partial x_j}\right)(x_0).$$

En lo que sigue, nos concentraremos en el caso $m = 1$, pues el caso general es idéntico considerando las funciones coordenadas $f_1, f_2, \ldots, f_m : \Omega \to \mathbb{R}$. Si $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función tal que todas sus derivadas parciales existen y son continuas en Ω, el vector gradiente de f es, en sí mismo, una función continua de $\mathbb{R}^n$ a $\mathbb{R}^n$, dada por

$$\begin{aligned} \nabla f : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}. \end{aligned}$$

Por lo tanto, podemos estudiar su diferenciabilidad.

Definición 2.5.2 (Matriz Hessiana)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función tal que todas sus derivadas parciales existen y son continuas en Ω, y sea $x_0 \in \Omega$. Definimos, cuando existe, la matriz Hessiana de f en x_0 como el Jacobiano de ∇f en x_0, esto es*

$$D^2 f(x_0) = D(\nabla f)(x_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{bmatrix}$$

Ejemplo 2.5.3. Para $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = x \sin(y) + y \cos(x),$$

su matriz hessiana está dada por

$$D^2 f(x, y) = \begin{bmatrix} -y\cos(x) & \cos(y) - \sin(x) \\ \cos(y) - \sin(x) & -x\sin(y) \end{bmatrix}$$

$\lozenge$

En el ejemplo anterior, es posible observar que

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \cos(y) - \sin(x) = \frac{\partial^2 f}{\partial y \partial x}(x, y).$$

Este hecho, que parece ser fortuito, no es tan así; gracias al siguiente resultado.

Teorema 2.5.4 (Teorema de Schwartz)**.** *Sea $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ y $x_0 \in \mathbb{R}^n$ tal que:*

1. *Todas las derivadas parciales de f existen y son continuas en Ω.*
2. *Todas las derivadas parciales $\frac{\partial f}{\partial x_j} : \Omega \to \mathbb{R}$ son* ***diferenciables*** *en Ω.*
3. *Todas las derivadas parciales de segundo orden son continuas en x_0.*

Entonces, para todo $i, j \in \{1, \ldots, n\}$,

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0).$$

Demostración. Siguiendo una estructura de demostración similar al del Teorema (**??**); asumamos $n = 2$ y $m = 1$ para comenzar. Consideremos la cantidad

$$o(h_1, h_2) = f(x_0 + h_1, y_0 + h_2) - f(x_0 + h_1, y_0) - f(x_0, y_0 + h_2) + f(x_0, y_0).$$

Si fijamos $h_2 \in \mathbb{R}$ y definimos la función

$$\varphi(t) = [f(x_0 + t, y_0 + h_2) - f(x_0, y_0 + h_2)] - [f(x_0 + t, y_0) - f(x_0, y_0)],$$

podemos aplicar el TVM para concluir que existe $t_1 \in (0, h_1)$ tal que

$$\varphi(h_1) - \varphi(0) = \varphi'(t_1)h_1,$$

lo que, reemplazando por lo que corresponde, implica que

$$o(h_1, h_2) = h_1 \left[\frac{\partial f}{\partial x}(x_0 + t_1, y_0 + h_2) - \frac{\partial f}{\partial x}(x_0 + t_1, y_0) \right].$$

Si usamos ahora el TVM para la función

$$\phi(t) = \frac{\partial f}{\partial x}(x_0 + t_1, y_0 + t) - \frac{\partial f}{\partial x}(x_0 + t_1, y_0),$$

encontraremos $t_2 \in (0, h_2)$ tal que

$$\frac{\partial f}{\partial x}(x_0 + t_1, y_0 + t_2) - \frac{\partial f}{\partial x}(x_0 + t_1, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0 + t_1, y_0 + t_2)h_2.$$

En consecuencia, tomando $(h_1, h_2) \to (0, 0)$, concluimos que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{o(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0).$$

Si hacemos lo mismo, pero fijando primero h_1 y luego h_2, concluimos que

$$\lim_{(h_1,h_2)\to(0,0)} \frac{o(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0),$$

y por unicidad del límite, se concluye lo deseado. En el caso de $n > 2$ la demostración es la misma, pues las únicas variables que se mueven en el caso general son x_j y x_k, por lo que se replica el desarrollo aquí planteado. Finalmente, para $m > 1$, el desarrollo se replica coordenada a coordenada para concluir. $\square$

Definición 2.5.5 (Funciones de Clase $\mathcal{C}^k$)**.** *Una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ se dice:*

1. *De clase $\mathcal{C}^1$ si todas sus derivadas parciales (de primer orden) existen y son continuas en Ω.*
2. *De clase $\mathcal{C}^k$ si es de clase $\mathcal{C}^{k-1}$ y todas sus derivadas parciales de orden k existen y son continuas en Ω.*

Con esta definición, podemos escribir un corolario del teorema anterior.

Corolario 2.5.6. *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^2$. Entonces, para todo $x \in \Omega$, el Hessiano de f en x, $D^2 f(x)$ es una matriz simétrica.*

2.5.1 Aproximaciones de Taylor

Cuando una función $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ es de clase $\mathcal{C}^2$, es posible aproximarla de mejor forma que sólo linealmente, usando los siguientes resultados (por ahora, sin demostración).

Teorema 2.5.7 (Taylor de Segundo Orden, Primera Versión)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^2$, y $x_0 \in \Omega$. Entonces,*

$$f(x_0 + h) = f(x_0) + Df(x_0)h + \frac{1}{2}h^T D^2 f(x_0)h + o(\|h\|^2),$$

donde el error $o(\|h\|^2)$ verifica

$$\frac{|o(\|h\|^2)|}{\|h\|^2} \to 0$$

cuando $h \to 0$.

Teorema 2.5.8 (Taylor de Segundo Orden, Segunda Versión)**.** *Sea $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^2$, y $x_0 \in \Omega$. Entonces, existe $z \in (x_0, x_0 + h)$ tal que*

$$f(x_0 + h) = f(x_0) + Df(x_0)h + \frac{1}{2}h^T D^2 f(z)h.$$

Estos teoremas, serán utilizados en demostraciones importantes del próximo capítulo del presente apunte.

2.6 Selección de Problemas

1. Encuentre la matriz jacobiana de la función $f : \mathbb{R}^2 \to \mathbb{R}^3$ dada por

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

2. Supongamos que estamos sobre el punto $(3, -5, 89)$ en un cerro cuya superficie está dada por la ecuación

$$z = 83 - x^2 - 6xy - 3y^2.$$

   Consideremos que en el eje Y apuntamos hacia arriba el norte, en el eje X apuntamos hacia la derecha el este y la distancia se mide en metros.

   a) Si nos movemos hacia el sur, ¿subimos o bajamos?, ¿con qué rapidez?

   b) ¿En qué dirección está el descenso más rápido?

   c) ¿En qué direcciones no hay ascenso ni descenso?

3. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x, y) = \begin{cases} (xy)^\alpha \cos\left(\dfrac{2\pi\alpha}{x^2 + y^2}\right) & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}.$$

   a) Muestre que si $\alpha > 0$, entonces f es continua en todo $\mathbb{R}^2$.

   b) Muestre que si $\alpha \leq 0$, entonces f no es continua en el origen.

c) Muestre que las derivadas direccionales de f en $(0,0)$ existen para todo $\alpha > \frac{1}{2}$, y además

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

d) Muestre que la función f es diferenciable en $(0,0)$, para todo $\alpha > \frac{1}{2}$.

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

$$f(x,y) = \begin{cases} \dfrac{\operatorname{sen}\left(\sqrt{x^2+y^2}\right)}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0) \\ \alpha & \text{si } (x,y) = (0,0) \end{cases}.$$

a) Encuentre el valor de $\alpha \in \mathbb{R}$ que hace que f sea una función continua.

b) Calcule las derivadas parciales de f para $(x,y) \neq (0,0)$.

c) Calcule las derivadas parciales de f para $(x,y) = (0,0)$.

5. Considere la función dada por

$$f(x,y) = \begin{cases} \dfrac{\ln(1+x^2+y^2) + x^3 + y^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ \alpha & \text{si } (x,y) = (0,0) \end{cases}$$

a) Encuentre el valor de α para que f sea una función continua en todo $\mathbb{R}^2$.

b) Calcule las derivadas parciales de f para $(x,y) \neq (0,0)$.

c) Calcule las derivadas parciales de f para $(x,y) = (0,0)$.

6. En economía, se dice que dos bienes son *sustitutos* si la demanda q_1 del primero crece cuando el precio p_2 del segundo crece y si la demanda q_2 del segundo crece cuando el precio p_1 del primero crece.

a) ¿Qué ejemplo conoce de bienes sustitutos?

b) Si dos bienes son sustitutos, ¿qué condiciones deben cumplir las derivadas parciales

$$\frac{\partial q_1}{\partial p_2} \text{ y } \frac{\partial q_2}{\partial p_1} \text{ ?}$$

c) Suponga que las funciones de demanda de dos bienes son

$$q_1 = 3000 + \frac{400}{p_1+3} + 50p_2, \quad q_2 = 2000 - 100p_1 + \frac{500}{p_2+4}.$$

¿Son sustitutos estos bienes?

7. De forma opuesta a la descrita en la pregunta anterior, dos bienes se dicen *complementarios* si la demanda q_1 del primero decrece cuando el precio p_2 del segundo crece y si la demanda q_2 del segundo decrece cuando el precio p_1 del primero aumenta.

a) ¿Qué bienes complementarios conoce?

b) Si dos bienes son complementarios, ¿qué condiciones deben cumplir las derivadas parciales

$$\frac{\partial q_1}{\partial p_2} \text{ y } \frac{\partial q_2}{\partial p_1} \text{ ?}$$

c) Suponga que las funciones de demanda de dos bienes son

$$q_1 = 2000 + \frac{400}{p_1+3} - 50p_2, \quad q_2 = 2000 - 100p_2 + \frac{500}{p_1+4}.$$

¿Son sustitutos estos bienes?

8. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ una función de clase $\mathcal{C}^1$ y sea $g : \mathbb{R}^2 \to \mathbb{R}^2$ una función definida a partir de f, como

$$g(u,v) = f(x(u,v),\ y(u,v),\ z(u,v)) = f(\cos(u) + \operatorname{sen}(v),\ \operatorname{sen}(u) + \cos(v),\ e^{u-v}).$$

a) Demuestre que g también es de clase $\mathcal{C}^1$.

b) Usando la regla de la cadena para jacobianos, y sabiendo además que

$$Df(1,1,1) = \begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 3 \end{bmatrix},$$

calcule la matriz jacobiana $Dg\left(\frac{\pi}{2}, \frac{\pi}{2}\right)$.

9. Encuentre el valor que debe tener $\lambda \in \mathbb{R}$ para que la matriz

$$A(x,y) = \begin{bmatrix} x & \lambda \\ 1 & y \end{bmatrix},$$

sea la matriz hessiana de una función f de clase $\mathcal{C}^2$. Luego, encuentre una función f tal que

$$D^2 f(x,y) = A(x,y).$$

10. Sea $u : \mathbb{R}^2 \to \mathbb{R}$ una función que satisface la siguiente ecuación en derivadas parciales

$$\frac{\partial u}{\partial t}(x,t) + u(x,t)\frac{\partial u}{\partial x}(x,t) = 0,$$

válida para todo $(x,y) \in \mathbb{R}^2$. Usando adecuadamente la regla de la cadena, encuentre una condición sobre $f : \mathbb{R} \to \mathbb{R}$ para que la función g definida por

$$g(t) = u(f(t), t)$$

sea constante como función de t.
Indicación: Calcule la derivada de g con respecto a t, y use su resultado para concluir.

11. Considere dos funciones $u(x,y)$ y $v(x,y)$ de clase $\mathcal{C}^2$, tales que

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

a) Muestre que u y v verifican

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

b) Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathcal{C}^2$ tal que

$$\frac{\partial^2 f}{\partial u^2} + \frac{\partial^2 f}{\partial v^2} = 1.$$

Definimos $g(x,y) = f(u(x,y), v(x,y))$, donde u y v son las funciones de la pregunta anterior. Pruebe que g satisface

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2.$$

c) **Verdadero o Falso:** En este contexto, si $u(x,y) = x^3 - 3xy^2$ y $v(x,y) = 3x^2y - y^3$, entonces

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 9(x^2 + y^2).$$

Justifique su respuesta, con argumentos matemáticos pertinentes.

12. Sea $f(u,v)$ una función dos veces diferenciable, definimos su *laplaciano* $\triangle f$ como la expresión

$$\triangle f(u,v) = \frac{\partial^2 f}{\partial u^2}(u,v) + \frac{\partial^2 f}{\partial v^2}(u,v).$$

El objetivo de este problema, será estudiar el laplaciano de un tipo particular de funciones.

Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función de clase $\mathcal{C}^2$; y definamos a su vez las funciones

$$u(x,y) = e^x \cos(y) \ , \quad v(x,y) = e^x \operatorname{sen}(y) \ .$$

Definimos además la función $g : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$g(x,y) = f(u(x,y), v(x,y)).$$

a) Pruebe que la función g cumple la ecuación

$$\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = e^{2x}\left(\frac{\partial^2 f}{\partial u^2} + \frac{\partial^2 f}{\partial v^2}\right).$$

b) Decimos que una función es armónica, si su laplaciano es igual a cero. Usando lo probado en la pregunta anterior, muestre que si f es una función armónica, entonces g también lo es.

13. Considere la función

$$\varphi(x,t) = A \operatorname{sen}(x + ct) + B \cos(x - ct),$$

donde A, B son constantes y $c > 0$ representa la velocidad de la luz en el vacío. Muestre que φ cumple la *ecuación de ondas*, dada por

$$\frac{\partial^2 \varphi}{\partial t^2} - c^2 \frac{\partial^2 \varphi}{\partial x^2} = 0.$$

Capítulo 3

Optimización en $\mathbb{R}^n$

3.1 Problemas de Optimización

Una de las aplicaciones más importantes del cálculo diferencial es la optimización. Esta es la rama de la matemática aplicada que estudia los procesos de decisión, y cómo tomar la *mejor decisión posible*. Todo problema de optimización cuenta con cuatro componentes:

1. Si es un problema de **maximización** o de **minimización**: La "mejor decisión posible" puede estar orientada a maximizar un beneficio, o a minimizar un costo.

2. Las **variables**: Esto corresponde a los elementos del problema que son controlados. En el contexto de este curso, corresponde a un vector $x \in \mathbb{R}^n$. Lo que se decide es el valor del vector x.

3. La **función objetivo**: El concepto de "mejor decisión" se cuantifica a través de una función $f : \mathbb{R}^n \to \mathbb{R}$. Para $x \in \mathbb{R}^n$ fijo, el valor $f(x)$ se interpreta como el valor de la decisión. Cuando el problema es de minimización, buscamos tomar la decisión x^* tal que $f(x^*)$ sea el menor valor posible. Cuando el problema es de maximización, buscamos tomar la decisión x^* tal que $f(x^*)$ sea el mayor valor posible.

4. Las **restricciones**: Normalmente, no podemos asignar cualquier valor (vectorial) a la variable x. En cambio, debemos respetar que el valor que asignamos a x pertenezca a un conjunto $X \subseteq \mathbb{R}^n$ de valores admisibles. El conjunto X se conoce como **conjunto factible**. La restricción corresponde a respetar la inclusión $x \in X$ al momento de decidir el valor de la variable x. El conjunto factible X se interpreta como lo "posible" en el problema de decisión.

Una vez identificados estos cuatro elementos, la forma estándar de escribir el problema de optimización asociado es:

$$\left\{\begin{array}{ll} \displaystyle\min_{x\in\mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{array}\right. \qquad \text{o bien,} \qquad \left\{\begin{array}{ll} \displaystyle\max_{x\in\mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{array}\right. \tag{3.1}$$

El problema de la izquierda en (3.1) corresponde a un problema de minimización, mientras que el problema de la derecha corresponde a un problema de maximización. El acrónimo *s.a.* significa *sujeto a*,enfatizando las restricciones del problema. La descripción de la nomenclatura formal está dada en la Figura 3.1.

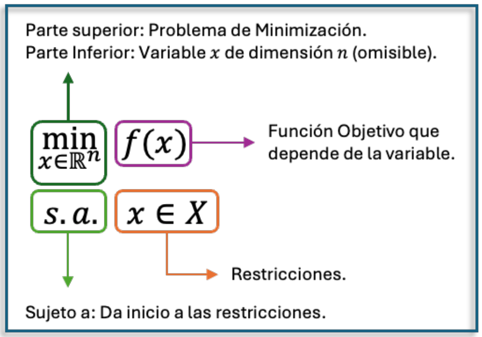


Figura 3.1: Esquema de problema de optimización

Nota: Problemas sin restricciones

Decimos que un problema de optimización es *sin restricciones* cuando el conjunto factible X coincide con todo el espacio, es decir, $X = \mathbb{R}^n$. En tal caso, escribimos simplemente

$$\min_{x\in\mathbb{R}^n} f(x) \qquad \text{o bien,} \qquad \max_{x\in\mathbb{R}^n} f(x), \tag{3.2}$$

según corresponda.

En muchos casos, es posible omitir el subtexto "$x \in \mathbb{R}^n$" donde se establece la variable y la dimensión. Eso se hace si no hay confusión. Es decir, podemos seguir el siguiente esquema:

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{cases} \xrightarrow{\text{omitir dimensión}} \begin{cases} \min\limits_{x} & f(x) \\ s.a. & x \in X. \end{cases} \xrightarrow{\text{omitir variable}} \begin{cases} \min & f(x) \\ s.a. & x \in X. \end{cases} \tag{3.3}$$

La última consideración que hay que tener en cuenta es que, en general, el conjunto factible X está dado por un conjunto de **ecuaciones e inecuaciones**. Es decir, en general

$$X = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} h_1(x) = 0, \dots, h_p(x) = 0, \\ g_1(x) \le 0, \dots, g_q(x) \le 0 \end{array} \right\} \tag{3.4}$$

donde $h_1, \dots, h_p$ y $g_1, \dots, g_q$ son funciones de $\mathbb{R}^n$ en $\mathbb{R}$. En estos casos, los problemas de optimización suelen escribirse poniendo las ecuaciones e inecuaciones directamente. Es decir, seguimos el siguiente esquema:

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{cases} \xrightarrow{\text{versión expandida}} \begin{cases} \min\limits_{x\in\mathbb{R}^n} & f(x) \\ s.a. & \begin{cases} h_1(x) = 0, \\ \vdots \\ h_p(x) = 0, \\ g_1(x) \le 0, \\ \vdots \\ g_q(x) \le 0. \end{cases} \end{cases} \tag{3.5}$$

Nota: Ecuaciones e inecuaciones vectoriales

Considerando las funciones vectoriales $h : \mathbb{R}^n \to \mathbb{R}^p$ dada por $h(x) = (h_1(x), \ldots, h_p(x))$, y $g : \mathbb{R}^n \to \mathbb{R}^p$ dada por $g(x) = (g_1(x), \ldots, g_q(x))$, podemos reescribir (3.4) como

$$X = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} h(x) = 0 \\ g(x) \leq 0 \end{array} \right\}, \tag{3.6}$$

donde en este caso, la inecuación $g(x) \leq 0$ se entiende como una inecuación puntual (es decir, $g(x) \leq 0$ si y solo si todas sus coordenadas son menores o iguales a cero). Esta forma de escribir nos permite usar una expansión parcial de los problemas de optimización.

$$\left\{ \begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{array} \right. \xrightarrow{\text{versión expandida parcial}} \left\{ \begin{array}{ll} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ s.a. & \left\{ \begin{array}{l} h(x) = 0, \\ g(x) \leq 0. \end{array} \right. \end{array} \right. \tag{3.7}$$

Si bien la forma (3.5) es mucho más común al momento de escribir un problema de optimización en la práctica, la forma (3.7) es muy común al momento de estudiar propiedades teóricas.

3.1.1 Algunos ejemplos

Antes de continuar, presentamos algunos ejemplos de problemas de optimización.

Ejemplo 3.1.1. Se desea encontrar el valor mínimo de la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x^2 + 3y^2$, en el conjunto

$$X = \left\{ (x, y) \in \mathbb{R}^2 \,|y = \cos(x), x \leq 2 \right\}$$

El problema de optimización asociado se puede escribir como

$$\left\{ \begin{array}{ll} \min\limits_{x,y} & x^2 + 3y^2 \\ s.a. & \begin{cases} y - \cos(x) = 0, \\ x \leq 2. \end{cases} \end{array} \right. \tag{3.8}$$

Si se desea encontrar el valor máximo de la función $f : \mathbb{R}^4 \to \mathbb{R}$ dada por $f(x, y) = x_1^2 - 3x_2^2 + 9x_3 - 27x_4$, en el conjunto

$$X = \left\{ x \in \mathbb{R}^4 \;\middle|\; \begin{array}{l} \|x\| \leq 1, \\ x_1 + x_2 = \sqrt{x_3 + x_4}, \\ x \geq 0 \end{array} \right\}.$$

El problema de optimización asociado se puede escribir como

$$\left\{ \begin{array}{ll} \max\limits_{x \in \mathbb{R}^4} & x_1^2 - 3x_2^2 + 9x_3 - 27x_4 \\ s.a. & \begin{cases} \|x\| \leq 1, \\ x_1 + x_2 = \sqrt{x_3 + x_4}, \\ x \geq 0. \end{cases} \end{array} \right. \tag{3.9}$$

◊

Atención

En este capítulo escribimos que las restricciones fueran de la forma $h(x) = 0$ o $g(x) \leq 0$. Sin embargo, todas las ecuaciones e inecuaciones se pueden reducir a esta forma. Concretamente:

- Una ecuación de la forma $a(x) = b(x)$ es equivalente a escribir $a(x) - b(x) = 0$.
- Una inecuación de la forma $a(x) \leq b(x)$ es equivalente a escribir $a(x) - b(x) \leq 0$.
- Una inecuación de la forma $c(x) \geq 0$ es equivalente a escribir $-c(x) \leq 0$.

Por lo tanto, admitiremos cualquier tipo de ecuación/inecuación dentro de las restricciones. Por ejemplo, como está escrito (3.9).

Ejemplo 3.1.2. Un problema clásico de optimización consiste en maximizar el área de una figura respetando restricciones geométricas, como por ejemplo que el perímetro sea constante. Podemos considerar un pentágono formado por un rectángulo inferior y un triángulo superior, como se muestra al lado izquierdo de Figura 3.2. Si se desea maximizar el área del pentágono, respetando que el perímetro

Figura 3.2: Pentágono formado por un rectángulo inferior y un triángulo superior.

tenga un valor P fijo, podemos proceder de la siguiente manera:

1. identificamos que el problema es de **maximización**.
2. identificamos **tres variables**: los lados x e y del rectángulo, y la altura h del triángulo.
3. La función objetivo es el área del pentágono que, en función de las variables elegidas, está dada por
$$f(x, y, h) = xy + \frac{xh}{2},$$
que corresponde a la suma de las áreas del rectángulo inferior y el triángulo superior.
4. Finalmente, las restricciones son la positividad de las variables ($x, y, h \geq 0$), y que el perímetro del pentágono sea igual a P, es decir,
$$x + 2y + 2\sqrt{\tfrac{x^2}{4} + h^2} = P.$$
Aquí, el tercer término de la suma del lado izquierdo corresponde a los lados del triángulo distintos de la base, calculados usando el teorema de Pitágoras para el triángulo rectángulo de catetos $\frac{x}{2}$ y h.

El problema de optimización resultante es

$$\begin{cases} \max\limits_{x,y,h} & xy + \frac{xh}{2} \\ s.a. & \begin{cases} x + 2y + 2\sqrt{\frac{x^2}{4} + h^2} = P, \\ x, y, h \geq 0. \end{cases} \end{cases} \tag{3.10}$$

$\diamondsuit$

Ejemplo 3.1.3. El problema de *despacho eléctrico* consiste en satisfacer la demanda de energía eléctrica a menor costo posible. Para esto, se puede considerar un modelo simplificado de la red de transmisión llamado *modelo uninodal.*
En un instante dado, consideramos la demanda acumulada de energía eléctrica, que denotamos D (medida en Mega-Watts [MW]). Suponemos que tenemos n productores de energía $P_1, \ldots, P_n$, donde el i-ésimo productor puede producir una cantidad q_i de potencia eléctrica (medida en Mega-Watts [MW]) e inyectarla a la red. El modelo uninodal consiste en suponer que la red de transmisión es un solo nodo, donde se inyectan las producciones $q_1, \ldots, q_n$ y se consume la demanda D.

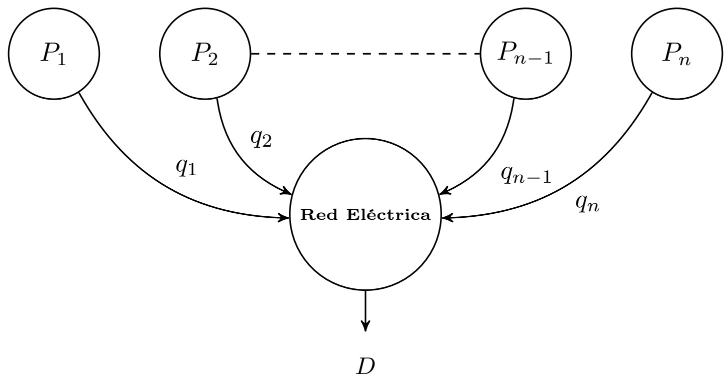


Figura 3.3: Modelo Uninodal para el despacho de energía eléctrica.

Cada productor tiene una función de costos $C_i : [0, +\infty) \to \mathbb{R}$ que determina su costo de producción: el costo de producir una potencia q_i está dado por $C_i(q_i)$. En muchos países (incluyendo Chile), existe un regulador central que debe decidir cuanta potencia produce cada productor, considerando la función de costos de cada uno, de manera de satisfacer la demanda a costo mínimo. Es decir, el regulador debe resolver el siguiente problema de optimización:

$$\left\{ \begin{array}{ll} \min\limits_{q \in \mathbb{R}^n} & \sum_{i=1}^n C_i(q_i) \\ s.a. & \begin{cases} \sum_{i=1}^n q_i = D \\ q_1, \ldots, q_n \geq 0. \end{cases} \end{array} \right. \tag{3.11}$$

En este problema de **minimización**, la **variable de decisión** es el vector $q = (q_1, \ldots, q_n)$ de potencias producidas, pues es lo que controla el agente. La **función objetivo** (a minimizar) es el costo total, dado por la suma de los costos. Es decir: $f(q) = \sum_{i=1}^n C_i(q_i)$. Las **restricciones** del problema son: 1) la positividad de las producciones $q_1, \ldots, q_n \geq 0$, pues no se puede producir una cantidad negativa de energía; y 2) Satisfacer la demanda, es decir, $\sum_{i=1}^n q_i = D$. ♢

3.2 Existencia de Soluciones

El objetivo de esta sección es definir lo que entenderemos por solución de un problema de optimización. Una vez que tengamos esta definición, también revisaremos la existencia de soluciones: es decir, bajo qué condiciones podemos asegurar que un problema de optimización tenga solución.

3.2.1 Noción de solución: óptimos globales y locales

Una vez planteado el problema de optimización de la forma (3.1), o en alguna de sus formas expandidas, necesitamos definir lo que será una solución del mismo. Para esto, consideramos la siguiente definición:

Definición 3.2.1. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función y $X \subseteq \mathbb{R}^n$ un conjunto. Decimos que un punto $\bar{x} \in X$ es*

1. ***mínimo global*** *de f en X si para todo $x \in X$, se tiene que $f(\bar{x}) \leq f(x)$, es decir, si*
$$\forall x \in X, f(\bar{x}) \leq f(x). \tag{3.12}$$
2. ***máximo global*** *de f en X si para todo $x \in X$, se tiene que $f(\bar{x}) \geq f(x)$, es decir, si*
$$\forall x \in X, f(\bar{x}) \geq f(x). \tag{3.13}$$
3. ***mínimo local*** *de f en X si existe $\delta > 0$, tal que para todo $x \in X \cap B(\bar{x}, \delta)$, se tiene que $f(\bar{x}) \leq f(x)$, es decir, si*
$$\exists \delta > 0, \forall x \in X \cap B(\bar{x}, \delta), f(\bar{x}) \leq f(x). \tag{3.14}$$
4. ***máximo local*** *de f en X si existe $\delta > 0$, tal que para todo $x \in X \cap B(\bar{x}, \delta)$, se tiene que $f(\bar{x}) \geq f(x)$, es decir, si*
$$\exists \delta > 0, \forall x \in X \cap B(\bar{x}, \delta), f(\bar{x}) \geq f(x). \tag{3.15}$$

La diferencia fundamental entre un mínimo global y un mínimo local es que un mínimo local $\bar{x}$ cumple la desigualdad $f(\bar{x}) \leq f(x)$ solamente "cerca de $\bar{x}$". Esto se ilustra en la Figura 3.4. En cierto sentido, podríamos catalogar los mínimos locales como "mínimos falsos" o "mínimos parciales".

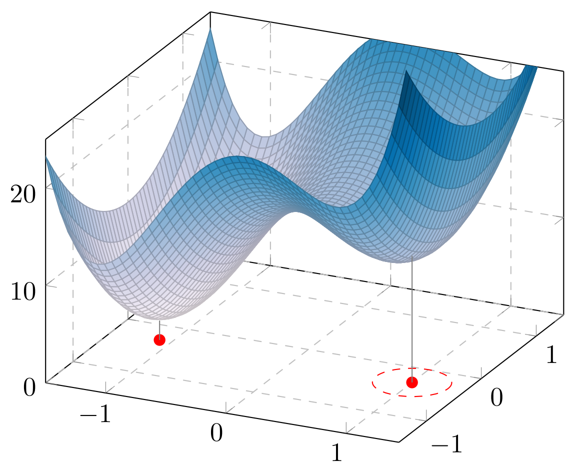


Figura 3.4: Mínimo global versus mínimo local: El punto de la derecha es mínimo local, pues es donde se alcanza el valor más pequeño dentro de la bola delimitada por la línea punteada roja. El punto de la izquierda es el mínimo global.

Otra definición importante es el valor de un problema de optimización. La idea es que, antes de definir lo que será una solución, podemos identificar el "mejor valor" que puede alcanzar la función objetivo. Este valor se define como el ínfimo (en caso de problemas de minimización) o el supremo (en caso de problemas de maximización) de los valores que toma la función objetivo f en el conjunto factible X.

Definición 3.2.2 (Valor de un problema)**.** *Considere el problema de optimización*

$$(P) = \begin{cases} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{cases}$$

Definimos el valor del problema (P) como

$$v(P) = \text{ínf}\{f(x) \ : \ x \in X\}, \tag{3.16}$$

*admitiendo la posibilidad de que $v(P) = -\infty$. Por convención, si $X = \emptyset$, decimos que el problema (P) es **infactible** y asignamos $v(P) = +\infty$. Análogamente, para el problema de optimización*

$$(P) = \begin{cases} \displaystyle\max_{x \in \mathbb{R}^n} & f(x) \\ s.a. & x \in X, \end{cases}$$

definimos el valor del problema (P) como

$$v(P) = \sup\{f(x) \ : \ x \in X\}, \tag{3.17}$$

*admitiendo la posibilidad de que $v(P) = +\infty$. Por convención, si $X = \emptyset$, decimos que el problema (P) es **infactible** y asignamos $v(P) = -\infty$.*

Con estos conceptos, definiremos una solución de un problema de optimización como un punto que alcanza el valor del problema. Es decir, un mínimo global o máximo global (según corresponda) de la función objetivo en el conjunto factible.

Definición 3.2.3 (Solución de un problema)**.** *Considere el problema de optimización*

$$(P) = \begin{cases} \displaystyle\min_{x \in \mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{cases}$$

Decimos que un punto $x^* \in \mathbb{R}^n$ es solución de (P) si $x^* \in X$ y $f(x^*) = v(P)$. Equivalentemente, $x^* \in \mathbb{R}^n$ es solución de (P) si y solamente si x^* es mínimo global de f en X.
Análogamente, considere el problema de optimización

$$(P) = \begin{cases} \displaystyle\max_{x \in \mathbb{R}^n} & f(x) \\ s.a. & x \in X. \end{cases}$$

Decimos que un punto $x^* \in \mathbb{R}^n$ es solución de (P) si $x^* \in X$ y $f(x^*) = v(P)$. Equivalentemente, $x^* \in \mathbb{R}^n$ es solución de (P) si y solamente si x^* es máximo global de f en X.

En la definición anterior, establecimos que las soluciones del problema de minimizar/maximizar f en el conjunto factible X corresponden a los mínimos/máximos globales de f en X. Sin embargo, dichas soluciones no tienen por qué ser únicas. Denotamos entonces los conjuntos de mínimos y máximos globales de la siguiente manera:

$$\text{argmin}_X f = \{x^* \in X \ : \ x^* \text{ es mínimo global de } f \text{ en } X\}. \tag{3.18}$$

$$\text{argmax}_X f = \{x^* \in X \ : \ x^* \text{ es máximo global de } f \text{ en } X\}. \tag{3.19}$$

Estos conjuntos pueden ser vacíos, cuando los respectivos problemas de optimización no tienen solución.

Nota: Alcanzar extremos

Para una función $f : \mathbb{R}^n \to \mathbb{R}$ y un conjunto X, se dice que

1. f **alcanza su mínimo** en X si el problema de minimizar f en X tiene solución.
2. f **alcanza su máximo** en X si el problema de maximizar f en X tiene solución.

En lo que sigue, concentraremos nuestro estudio en **problemas de minimización**. La razón es que el problema de maximizar una función f sobre un conjunto X se puede resolver minimizando la función $-f$ sobre el mismo conjunto. Esto se formaliza en la siguiente proposición.

Proposición 3.2.4. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función, X un conjunto, y sea $\bar{x} \in X$. Se tiene que*

1. *El punto x es un máximo local de f en X si y solo si x es un mínimo local de $-f$ en X.*
2. *El punto x es un máximo global de f en X si y solo si x es un mínimo global de $-f$ en X.*

En particular, $\operatorname{argmax}_X f = \operatorname{argmin}_X(-f)$.

Demostración. Mostremos primero el caso de óptimo local. Aquí, podemos escribir

$$\begin{aligned}
x \text{ es máximo local de } f \text{ en } X &\iff \exists \delta > 0, \forall y \in B(x,\delta) \cap X, f(x) \geq f(y) \\
&\iff \exists \delta > 0, \forall y \in B(x,\delta) \cap X, -f(x) \leq -f(y) \\
&\iff x \text{ es mínimo local de } f \text{ en } X.
\end{aligned}$$

Para el caso de óptimo global, el desarrollo es análogo. En efecto, podemos escribir

$$\begin{aligned}
x \text{ es máximo global de } f \text{ en } X &\iff \forall y \in X, f(x) \geq f(y) \\
&\iff \forall y \in X, -f(x) \leq -f(y) \\
&\iff x \text{ es mínimo global de } f \text{ en } X.
\end{aligned}$$

Como conclusión del segundo desarrollo, podemos escribir

$$\begin{aligned}
\operatorname{argmax}_X f &= \{x \in X \ : \ x \text{ es máximo global de } f \text{ en } X\} \\
&= \{x \in X \ : \ x \text{ es mínimo global de } -f \text{ en } X\} \\
&= \operatorname{argmin}_X(-f).
\end{aligned}$$

Esto concluye la demostración. □

3.2.2 Teorema de Weierstrass

Para asegurar la existencia de soluciones de un problema de optimización, es natural que necesitemos condiciones sobre los dos elementos que definen el problema: la función objetivo y el conjunto factible.

Definición 3.2.5. *Un conjunto $X \subseteq \mathbb{R}^n$ se dice acotado, si existe $M > 0$, lo suficientemente grande tal que $X \subset \overline{B}(0, M)$. Es decir, si*

$$\exists M > 0, \forall x \in X, \|x\| \leq M. \tag{3.20}$$

El siguiente teorema, llamado *Teorema de Valores Extremos de Weiestrass*, es uno de los resultados fundamentales de la optimización, que nos entrega un contexto bastante amplio para asegurar existencia de soluciones. Muchos de los desarrollos modernos de existencia de soluciones buscan aplicar de una u otra manera este teorema, o alguna variante del mismo.

Teorema 3.2.6 (Teorema de Valores Extremos de Weierstrass)**.** *Sea $X \subseteq \mathbb{R}^n$ un conjunto no-vacío, **cerrado y acotado**, y sea $f : \mathbb{R}^n \to \mathbb{R}$ una función continua en X. Entonces, f alcanza su máximo y su mínimo en f.*

Demostración. Mostraremos que f alcanza su mínimo. La demostración para el máximo se deduce de reemplazar f por $-f$ y replicar el análisis. La idea es ir dividiendo X en hipercubos cada vez más pequeños, y siempre mantener el cubo donde el valor del problema de minimizar f se mantenga. Los vértices de estos hipercubos generarán sucesiones convergentes cuyo límite será la solución del problema de optimización.

Demostración en $\mathbb{R}$:

Estudiaremos primero la demostración para el caso $n = 1$, es decir, con $X \subseteq \mathbb{R}$.

Paso 0 - Acotamiento: Como X es cerrado y acotado, existen valores $a_0, b_0 \in \mathbb{R}$ tales que

$$X \subset I_0 = \{x \in \mathbb{R} : a_0 \leq x \leq b_0\}.$$

Denotemos por $B_0 = X$ y a v_0 el valor del problema $\min_{x \in B_0} f(x)$.

Paso 1 - Primera división: Tomemos el punto medio $m_0 = \frac{a_0+b_0}{2}$ y definamos los conjuntos $B_{0,1} = [a_0, m_0]$, $B_{0,2} = [m_0, b_0]$, y

$$X_{0,1} = X_0 \cap [a_0, m_0],$$
$$X_{0,2} = X_0 \cap [m_0, b_0].$$

Podemos definir los valores $v_{0,1} = \inf\{f(x) : x \in X_{0,1}\}$ y $v_{0,2} = \inf\{f(x) : x \in X_{0,2}\}$. Como $X_0 = X_{0,1} \cup X_{0,2}$, se tiene que

$$v_0 = \inf\{f(x) : x \in X_0\} = \min\{v_{0,1}, v_{0,2}\}.$$

Elijamos $i \in \{1, 2\}$ tal que $v_0 = v_{0,i}$, y definamos $X_1 = X_{0,i}$ y $B_1 = B_{0,i}$. Definimos $a_1 = \min B_1$ y $b_1 = \max B_1$. La construcción se ilustra en la Figura 3.5.

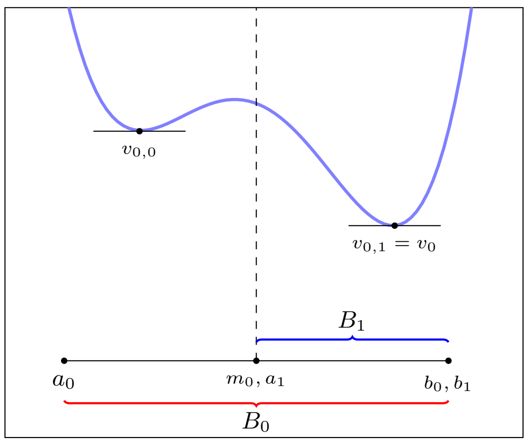


Figura 3.5: Ilustración de primera división. En este caso, se toma $B_1 = B_{0,1}$ y se descarta $B_{0,0}$.

En la construcción anterior, o bien reemplazamos a_0 por m_0 (cuando $B_1 = B_{0,2}$), o bien reemplazamos b_0 por m_0 (cuando $B_1 = B_{0,1}$). Hemos construido un conjunto X_1 y términos a_1, b_1 tales que

- $v_0 = \inf\{f(x) : x \in X_1\}$,
- $X_1 \subset [a_1, b_1]$ $(= B_1)$,
- $a_0 \leq a_1 \leq b_1 \leq b_0$, y
- $b_1 - a_1 = (b_0 - a_0)/2$.

Paso 2 - División iterativa: La idea es repetir lo que hicimos en el paso 1 infinitas veces. Esta construcción la extenderemos por inducción. Supongamos que hemos construido una secuencia de conjuntos $(X_j)_{j=0}^k$, y los términos $(a_j, b_j)_{j=1}^k$, cumpliendo que

1. $X_k \subset X_{k-1} \subset \cdots X_1 \subset X_0$,
2. $v_0 = \inf\{f(x) : x \in X_k\}$,

3. $X_k \subset [a_k, b_k]$,

4. $a_0 \leq a_1 \leq \cdots a_k \leq b_k \leq \cdots \leq b_1 \leq b_0$,

5. $b_j - a_j = (b_0 - a_0)/2^j$, para todo $j \in \{0, \ldots, k\}$.

Repitiendo la primera división reemplazando X_0 y B_0 por X_k y $B_k = [a_k, b_k]$, podemos construir X_{k+1} y $a_{k+1}, b_{k+1} \in \mathbb{R}$ que cumplen que

- $\inf\{f(x) : x \in X_{k+1}\} = \inf\{f(x) : x \in X_k\}$,
- $X_{k+1} \subset [a_{k+1}, b_{k+1}]$,
- $a_k \leq a_{k+1} \leq b_{k+1} \leq b_k$, y
- $b_{k+1} - a_{k+1} = (b_k - a_k)/2$.

Se tiene entonces que la secuencia hasta $k+1$ cumple las propiedades 1-5 también.

Paso 3 - conclusión: Repitiendo la construcción infinitas veces, tendremos una secuencia de conjuntos $(X_k)_{k \in \mathbb{N}}$ y dos sucesiones $(a_k)_{k \in \mathbb{N}}, (b_k)_{k \in \mathbb{N}}$ tales que

- $X_{k+1} \subset X_k$ para todo $k \in \mathbb{N}$,
- $v_0 = \inf\{f(x) : x \in X_k\}$ para todo $k \in \mathbb{N}$,
- $X_k \subset [a_k, b_k]$,
- (a_k) es creciente y (b_k) es decreciente,
- $b_k - a_k = (b_0 - a_0)/2^k$, para todo $k \in \mathbb{N}$.

Como la sucesión (a_k) es creciente y acotada ($a_0 \leq a_k \leq b_0$ para todo k), necesariamente converge a su supremo. Es decir,

$$\lim a_k = \sup\{a_k : k \in \mathbb{N}\} = \bar{a} \in \mathbb{R}.$$

Análogamente, como la sucesión (b_k) es decreciente y acotada ($a_0 \leq b_k \leq b_0$ para todo k), necesariamente converge a su ínfimo. Es decir,

$$\lim b_k = \sup\{b_k : k \in \mathbb{N}\} = \bar{b} \in \mathbb{R}.$$

Luego, tenemos que

$$\bar{b} - \bar{a} = \lim_k (b_k - a_k) = \lim_k \frac{b_0 - a_0}{2^k} = 0.$$

Por lo tanto, $\bar{a} = \bar{b}$. Denotemos el valor común por $\bar{x}$.
Ahora, usando que $\inf\{f(x) : x \in X_k\} = v_0$ para todo $k \in \mathbb{N}$ podemos elegir una sucesión $(x_k)_k$ con $x_k \in X_k$ y tal que

$$\begin{cases} f(x_k) \leq v_0 + \frac{1}{k}, & \forall k \in \mathbb{N}, \quad \text{si } v_0 \in \mathbb{R}, \\ f(x_k) \leq -k, & \forall k \in \mathbb{N}, \quad \text{si } v_0 = -\infty. \end{cases}$$

Luego, tenemos que

(i) Usando que $X_k \subset [a_k, b_k]$, tenemos que $a_k \leq x_k \leq b_k$ para todo $k \in \mathbb{N}$.

(ii) Usando que $X_k \subseteq X_0 = X$, tenemos que $(x_k)_{k \in \mathbb{N}} \subset X$.

(iii) Por construcción, $\lim f(x_k) = v_0$.

Por sándwich, (i) nos dice que $x_k \to \bar{x}$. Como X es cerrado, tenemos que $\bar{x} \in X$. Finalmente, por continuidad de f, tenemos que

$$v_0 = \lim f(x_k) = f(\bar{x}).$$

Por lo tanto v_0 no puede ser $-\infty$ y luego $v_0 \in \mathbb{R}$ y $f(\bar{x}) = v_0$. Así, f alcanza su mínimo en $\bar{x} \in X$.

Demostración en $\mathbb{R}^n$:

La demostración del caso general sigue la misma idea, excepto que ahora tenemos que dividir en hipercubos en vez de intervalos. Denotemos nuevamente $v_0 = \inf_{x\in X} f(x)$, que nuevamente podría ser $-\infty$.

Paso 0 - Acotamiento: Como X es cerrado y acotado, existen vectores $a^0 = (a_1^0, \dots, a_n^0), b^0 = (b_1^0, \dots, b_0^n) \in \mathbb{R}^n$ tales que

$$X \subset C_0 = \{x \in \mathbb{R} \,:\, a_i^0 \le x_i \le b_i^0,\ i = 1, \dots, n\}.$$

Denotemos por $X_0 = X$ y a v_0 el valor del problema $\min_{x\in X_0} f(x)$. Luego, existe $x \in X_0$, tal que

$$f(x_0) \le \begin{cases} v_0 + 1 & \text{si } v_0 \in \mathbb{R}, \\ 0 & \text{si } v_0 = -\infty. \end{cases}$$

Paso 1 - Primera división: Consideramos el vector

$$m^0 = \left(\frac{a_1^0 + b_1^0}{2}, \dots, \frac{a_n^0 + b_n^0}{2}\right).$$

Usando las coordenadas de m_0, podemos dividir C_0 en 2^n partes, eligiendo la parte $[a_i^0, m_i^0]$ o la parte $[m_i^0, b_i^0]$ para cada coordenada $i = 1, \dots, n$. Denotaremos estas subdivisiones como $C_{0,1}$ hasta $C_{0,2^n}$. Igual que en la demostración para $\mathbb{R}$, existe una de estas divisiones que preserva el ínfimo, es decir, que existe $k \in \{1, \dots, 2^n\}$ tal que el conjunto $X_1 = X \cap C_{0,k}$ es no-vacío y donde

$$\inf_{x\in X_1} f(x) = v_0.$$

Luego, existe $x_1 \in X_1$ tal que

$$f(x_1) \le \begin{cases} v_0 + 1/2 & \text{si } v_0 \in \mathbb{R}, \\ -1 & \text{si } v_0 = -\infty. \end{cases}$$

Denotamos $C_1 = C_{0,k}$ y tomamos a^1, b^1 como los vértices asociados a la diagonal principal de C_1. Es decir, $a^1, b^1 \in \mathbb{R}^n$ son tales que

$$C_1 = \{x \in R^n \,:\, a_i^1 \le x_i \le b_i^1,\ i = 1, \dots, n\}.$$

En efecto, si para la coordenada i elegimos el intervalo $[a_i^0, m_i^0]$, entonces $a_i^1 = a_i^0$ y $b_i^1 = m_i^0$. Por el contrario, si elegimos el intervalo $[m_i^0, b_i^0]$, entonces $a_i^1 = m_i^0$ y $b_i^1 = b_i^0$. En cualquier caso, se tiene que

$$\|b^1 - a^1\| = \frac{1}{2}\|b^0 - a^0\|.$$

La Figura 3.6 representa los pasos 0 y 1 en el caso $n = 2$.

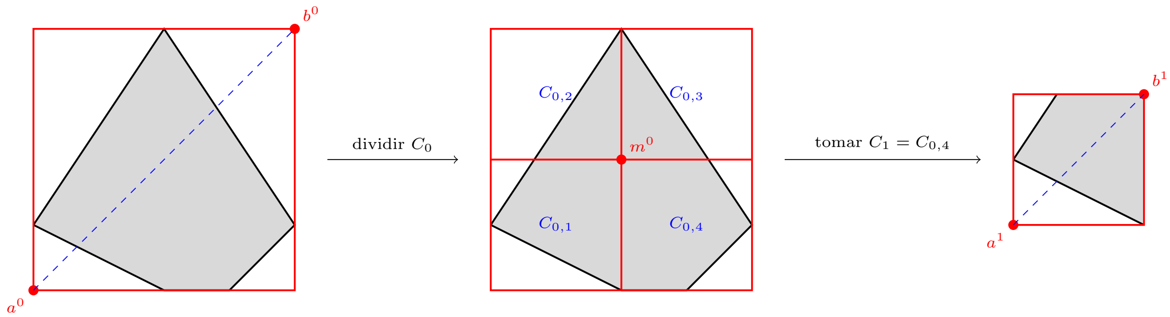


Figura 3.6: División en 4 partes. Se selecciona $C_1 = C_{0,4}$. La distancia de a^1 y b^1 se reduce a la mitad con respecto a la distancia entre a^0 y b^0.

Paso 2 - División iterativa: Podemos repetir el proceso, generando una secuencia de conjuntos $\{X_k : k \in \mathbb{N}\}$, y tres secuencias $(x_k)_{k\in\mathbb{N}}$ $(a^k)_{k\in\mathbb{N}}$ y $(b^k)_{k\in\mathbb{N}}$ de vectores en $\mathbb{R}^n$ tales que

- $X_k \subset X$ no-vacío, y $\inf_{x\in X_k} f(x) = v_0$.
- $x_k \in X_k$ y $f(x_k) \leq \begin{cases} v_0 + 1/(k+1) & \text{si } v_0 \in \mathbb{R}, \\ -k & \text{si } v_0 = -\infty. \end{cases}$
- $X_k \subset C_k = \{x : a_i^k \leq x_i \leq b_i^k,\ i = 1, \dots, n\}$.
- $\|b^k - a^k\| \leq \frac{1}{2}\|b^{k-1} - a^{k-1}\|$.

Paso 3 - conclusión: La construcción anterior tiene como propiedad que si fijamos una coordenada $i \in [n]$, tenemos que

- La sucesión coordenada $(a_i^k)_{k\in\mathbb{N}}$ es creciente y acotada. Por lo tanto,
$$a_i^k \xrightarrow{k\to\infty} \bar{a}_i = \sup\{a_i^k : k \in \mathbb{N}\}.$$
- La sucesión coordenada $(b_i^k)_{k\in\mathbb{N}}$ es decreciente y acotada. Por lo tanto,
$$b_i^k \xrightarrow{k\to\infty} \bar{b}_i = \inf\{b_i^k : k \in \mathbb{N}\}.$$

Concluimos entonces que $a^k \to \bar{a} = (\bar{a}_1, \dots, \bar{a}_n)$ y que $b^k \to \bar{b} = (\bar{b}_1, \dots, \bar{b}_n)$. Más aún,
$$\|\bar{b} - \bar{a}\| = \lim_k \|b^k - a^k\| \leq \lim_k \frac{\|b^0 - a^0\|}{2^k} = 0.$$
Por lo tanto, $\bar{a} = \bar{b}$. Finalmente, como $a_i^k \leq x_{k,i} \leq b_i^k$ para todo $i \in [n]$ concluimos, usando el teorema del Sándwich para cada coordenada (ver Proposición 1.2.13), que $x_k \to \bar{x}$ y que $\bar{x} = \bar{a} = \bar{b}$.
Como X es cerrado y $(x_k) \subset X$, tenemos que $\bar{x} \in X$. Luego, como f es continua, tenemos que
$$v_0 \leq f(\bar{x}) = \lim_k f(x_k) \begin{cases} \leq \lim_k \left(v_0 + \frac{1}{k+1}\right) = v_0 & \text{si } v_0 \in \mathbb{R}, \\ \leq \lim_k -k = -\infty = v_0 & \text{si } v_0 = -\infty. \end{cases}$$
Concluimos que en ambos casos $f(\bar{x}) = v_0$ y por lo tanto v_0 no puede ser $-\infty$. Así, f alcanza su mínimo en $\bar{x}$, lo que concluye la demostración. $\square$

Una aplicación del Teorema de Weiestrass es la existencia de mínimos para funciones coercitivas.

Definición 3.2.7 (Funciones coercitivas)**.** *Una función $f : \mathbb{R}^n \to \mathbb{R}$ se dice coercitiva si y solo si para toda sucesión $(x_k)_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$, se tiene que*

$$\|x_k\| \xrightarrow{k\to\infty} +\infty \implies f(x_k) \xrightarrow{k\to\infty} +\infty. \tag{3.21}$$

En tal caso, escribiremos simplemente que $\lim_{\|x\|\to+\infty} f(x) = +\infty$.

Nota: límites a infinito

Recordemos que para una función $f : \mathbb{R} \to \mathbb{R}$, los límites de valor infinito estaban bien definidos. En particular,

$$\lim_{t\to+\infty} f(t) = +\infty \iff \forall M > 0, \exists t_M > 0, \forall t \geq t_M, f(t) \geq M.$$

Similarmente, para una sucesión $(r_k)_k \subseteq \mathbb{R}$, definimos el límite de valor $+\infty$ como

$$\lim_{k\to\infty} r_k = +\infty \iff \forall M > 0, \exists k_0 > 0, \forall k \geq k_0, r_k \geq M.$$

Esto es lo que se aplica en la Definición 3.2.7.

Intuitivamente, las funciones coercitivas son aquellas que crecen a infinito en la medida que se alejan del origen. Con esto en mente, si una función coercitiva alcanza su mínimo, el mínimo global no puede estar muy lejos del origen. Con esta intuición, podemos plantear la siguiente proposición.

Teorema 3.2.8. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función coercitiva y continua, y sea X un conjunto cerrado no-vacío. Entonces, f alcanza su mínimo en X.*

Demostración. Como X es no-vacío, tomemos $x_0 \in X$ y denotemos $\alpha = f(x_0)$. Se tiene entonces que el conjunto

$$A := \{x \in X \ : \ f(x) \leq \alpha\}$$

es no-vacio, pues al menos contiene a x_0. Veamos primero que A es cerrado. Sea (x_k) una sucesión en A con $x_k \to \bar{x}$. Como X es cerrado y $A \subset X$, se tiene que $\bar{x} \in X$. Por otro lado, como f es continua, tenemos que

$$f(\bar{x}) = \lim_{k\to\infty} f(x_k) \leq \lim_{k\to\infty} \alpha = \alpha.$$

Por lo tanto, $\bar{x} \in A$. Como (x_k) es una sucesión convergente arbitraria, la Proposición **??** nos asegura que A es cerrado.
Veamos ahora que A es acotado. Razonando por contradicción, supongamos que no lo fuera. Entonces, para todo $k \in \mathbb{N}$, debe existir un punto $x_n \in A$ tal que $\|x_k\| > k$. Esto se obtiene aplicando la negación de (3.20) a $M = k$. Luego, $\|x_k\| \to +\infty$ y como f es coercitiva, $f(x_k) \to +\infty$. Sin embargo, como $(x_k) \subseteq A$, $f(x_k) \leq \alpha$, lo cual es una contradicción. Concluimos entonces que A es acotado.
Aplicando el Teorema 3.2.6, tenemos que el problema de optimización

$$(P) = \begin{cases} \min_{x\in\mathbb{R}^n} & f(x) \\ s.a. & x \in A, \end{cases}$$

tiene solución. Sea $x^* \in A$ una solución de (P). Entonces, se tiene que para todo $x \in A$, se cumple que $f(x^*) \leq f(x)$. Más aún, considerando el resto de puntos de X, es decir, $X \setminus A$, se cumple que

$$\forall y \in X \setminus A, f(x^*) \leq \alpha < f(y).$$

Es decir,

$$\forall x \in X = A \cup X \setminus A, f(x^*) \leq f(x).$$

Concluimos entonces que x^* es mínimo global de f en X, lo que concluye la demostración. □

3.3 Optimización sin restricciones

En esta sección nos enfocaremos en problemas de optimización sin restricciones, es decir, problemas donde el conjunto factible es todo el espacio $X = \mathbb{R}^n$.
Una pregunta natural es cómo encontrar máximos y mínimos de una función f. Esta no es una pregunta fácil de responder y existe todo un desarrollo de las matemáticas n torno a esta pregunta, tanto desde el punto de vista de cálculo teórico, como desde la perspectiva algorítmica. En este curso, nos concentraremos en una de las formas fundamentales de atacar el problema: reducirlo a resolver un sistema de ecuaciones.

Teorema 3.3.1 (Regla de Fermat)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable, y sea $x \in \mathbb{R}^n$. Si x es un mínimo local o un máximo local, entonces $\nabla f(x) = 0$.*

Demostración. Supongamos que $x \in \mathbb{R}^n$ es un mínimo local y sea $\delta > 0$ tal que para todo $y \in B(x, \delta)$ se cumple que $f(x) \leq f(y)$.
Supongamos que $\nabla f(x) \neq 0$ y definamos $d = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$. Como f es diferenciable, sabemos que

$$-\|\nabla f(x)\| = \langle \nabla f(x), d\rangle = f'(x; d) = \lim_{t\to 0} \frac{f(x + td) - f(x)}{t}.$$

Ahora, consideremos una secuencia $t_k = \frac{\delta}{2k} \in (0, \delta)$. Tenemos que $t_k \to 0$ y que $x + t_k d \in B(x, \delta)$ para todo $k \in \mathbb{N}$. En particular, $f(x + t_k d) - f(x) \geq 0$ para todo $k \in \mathbb{N}$. Tenemos entonces que

$$0 > -\|\nabla f(x)\| = \lim_{k\to\infty} \frac{f(x + t_k d) - f(x)}{t_k} \geq 0,$$

lo cual es una contradicción. Por lo tanto $\nabla f(x) = 0$. Si x es un máximo local, entonces usando la Proposición 3.2.4, tenemos que x es un mínimo local de $-f$. Entonces, usando el desarrollo anterior, tenemos que $\nabla(-f)(x) = 0$. La demostración termina notando que $\nabla f(x) = -\nabla(-f)(x)$ y por lo tanto $\nabla f(x) = 0$. □

Todo mínimo/máximo local cumple que $\nabla f(x) = 0$, pero puede haber puntos que no son ni máximos ni mínimos locales y que también cumplan $\nabla f(x) = 0$. Un ejemplo es el caso del punto $(0, 0)$ en la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = x^2 - y^2$ (ver Figura 2.2 del capítulo 2). Esto motiva la siguiente definición.

Definición 3.3.2 (Puntos críticos)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. Decimos que $x \in \mathbb{R}^n$ es **punto crítico** de f si $\nabla f(x) = 0$.*

La regla de Fermat y la noción de puntos críticos nos permiten establecer una metodología para resolver problemas de optimización. Consideremos entonces el problema

$$(P) = \min_{x\in\mathbb{R}^n} f(x),$$

donde f es diferenciable. Para resolver (P), podemos proceder de la siguiente manera:

1. Mostrar que f tiene solución (por ejemplo, usando la Proposición 3.2.8).
2. Encontrar todos los puntos críticos de f, resolviendo la ecuación $\nabla f(x) = 0$.
3. Dentro de los puntos críticos que encontramos, elegimos x^* tal que $f(x^*)$ tenga el valor más pequeño posible.
4. Todos los mínimos globales son mínimos locales, y por lo tanto son puntos críticos gracias al Teorema 3.3.1. Así que, el punto x^* escogido en el paso anterior tiene que ser el mínimo global.

Atención

El método anterior solo funciona si el paso 1 se verifica, es decir, si el problema (P) tiene solución. Si este no es el caso, ninguno de los puntos críticos nos servirá. Por ejemplo, consideremos nuevamente la función $f(x,y) = x^2 - y^2$ (ver Figura 2.2 del capítulo 2). Para esta función, tenemos que

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} = 0 \iff x = y = 0,$$

es decir, $(0,0)$ es el único punto crítico de f. Sin embargo, $(0,0)$ no es ni máximo ni mínimo global, pues f no alcanza su máximo ni su mínimo en $\mathbb{R}^n$. Peor aún, $(0,0)$ tampoco es máximo local ni mínimo local.

Aunque el método recién descrito nos permite encontrar soluciones en caso de existir, no nos entrega necesariamente como es la naturaleza de los puntos críticos encontrados. ¿Son máximos locales? ¿Son mínimos locales? o ¿Son puntos críticos que no son óptimos locales? Para responder estas preguntas, en el caso de las funciones que admitan segunda derivada, podemos estudiar la matriz Hessiana.

Nota: Matrices definidas positivas

Recordemos que una matriz cuadrada **simétrica** $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ se dice

1. **Semidefinida positiva** si para todo $x \in \mathbb{R}^n$, $x^T Ax \geq 0$.
2. **Definida positiva** si para todo $x \in \mathbb{R}^n \setminus \{0\}$, $x^T Ax > 0$.
3. **Semidefinida negativa** si para todo $x \in \mathbb{R}^n$, $x^T Ax \leq 0$.
4. **Definida negativa** si para todo $x \in \mathbb{R}^n \setminus \{0\}$, $x^T Ax < 0$.

Estas propiedades se pueden caracterizar usando los valores propios de la matriz: Para $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ simétrica se tiene que

$$\begin{aligned}
A \text{ es semidefinida positiva} &\iff \text{Todos sus valores propios son mayores o iguales a cero.}\\
A \text{ es definida positiva} &\iff \text{Todos sus valores propios son mayores estrictos a cero.}\\
A \text{ es semidefinida negativa} &\iff \text{Todos sus valores propios son menores o iguales a cero.}\\
A \text{ es definida negativa} &\iff \text{Todos sus valores propios son menores estrictos a cero.}
\end{aligned}$$

Proposición 3.3.3 (Criterio de segundo orden). *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^2$, y sea $x \in \mathbb{R}^n$ un punto crítico de f (es decir, tal que $\nabla f(x) = 0$). Entonces, se tiene que*

1. *Si $D^2 f(x)$ es definida positiva, entonces x es un mínimo local.*
2. *Si $D^2 f(x)$ es definida negativa, entonces x es un máximo local.*
3. *Si $D^2 f(x)$ tiene al menos un valor propio estrictamente positivo y al menos otro estrictamente negativo, entonces x es un punto silla (ni máximo local, ni mínimo local).*

Demostración. Demostremos la primera afirmación. Recordemos que si f es de clase $\mathcal{C}^2$, entonces $D^2 f(x)$ es una matriz simétrica. Usando el teorema de Taylor (ver Teorema 2.5.7), sabemos que

$$f(x+h) = f(x) + Df(x)h + \frac{1}{2}\langle h, D^2 f(x)h\rangle + o(\|h\|^2).$$

Definamos la función $g(d) = \frac{1}{2}\langle d, D^2 f(x)d\rangle$. Claramente esta función es continua, por lo tanto, el problema de optimización

$$(P) = \begin{cases} \min\limits_{d\in\mathbb{R}^n} & g(d) \\ s.a. & \|d\| = 1, \end{cases}$$

tiene solución. Sea d^* una solución de (P). Como $\|d^*\| = 1$, tenemos que $d^* \neq 0$ y por lo tanto, usando que $D^2f(x)$ es definida positiva, tenemos que $v(P) = g(d^*) > 0$. Denotemos por $\alpha = v(P)$. Recordando la definición de funciones o-chica (ver los comentarios bajo la Definición **??**), tenemos que existe $\delta > 0$ tal que

$$\forall h \in B(0, \delta), \frac{|o(\|h\|^2)|}{\|h\|^2} \leq \frac{\alpha}{2}.$$

Con esto en mente, tomando $y \in B(x, \delta)$ y notando que $h = y - x \in B(0, \delta)$, podemos escribir

$$\begin{aligned}
f(y) - f(x) &= Df(x)h + \frac{1}{2}\langle h, D^2f(x)h\rangle + o(\|h\|^2) \\
&= \frac{1}{2}\langle h, D^2f(x)h\rangle + o(\|h\|^2) \\
&= \|h\|^2 \left(\frac{1}{2}\left\langle \frac{h}{\|h\|}, D^2f(x)\frac{h}{\|h\|}\right\rangle + \frac{o(\|h\|^2)}{\|h\|^2}\right) \\
&= \|h\|^2 \left(g\left(\frac{h}{\|h\|}\right) + \frac{o(\|h\|^2)}{\|h\|^2}\right) \\
&\geq \|h\|^2 \left(\alpha - \frac{|o(\|h\|^2)|}{\|h\|^2}\right) \\
&\geq \|h\|^2 \left(\alpha - \frac{\alpha}{2}\right) = \frac{\|h\|^2\alpha}{2} > 0.
\end{aligned}$$

Como $y \in B(x, \delta)$ es arbitrario, concluimos que x es mínimo local de f.
La segunda afirmación se deduce de la primera notando que $D^2(-f)(x) = -D^2f(x)$, y por lo tanto, si $D^2f(x)$ es definida negativa, entonces $D^2(-f)(x)$ es definida positiva. Así, x será minimo local de $-f$ y por lo tanto, máximo local de f.
Para la tercera afirmación, consideremos v^+ un vector propio unitario (es decir, con norma 1) de $D^2f(x)$ asociado a un valor propio $\lambda^+ > 0$, y v^- un vector propio unitario de $D^2f(x)$ asociado a un valor propio $\lambda^- < 0$. Recordar que esto significa que $D^2f(x)v^+ = \lambda^+v^+$ y que $D^2f(x)v^- = \lambda^-v^-$.
Usando nuevamente el Teorema de Taylor, tenemos que

$$\begin{aligned}
f(x + tv^+) - f(x) &= +tDf(x)v^+ + \frac{1}{2}\langle tv^+, D^2f(x)tv^+\rangle + o(t^2) \\
&= \frac{t^2}{2}\lambda^+ o(t^2) \\
&= t^2\left(\frac{\lambda^+}{2} - \frac{|o(t^2)|}{t^2}\right).
\end{aligned}$$

Tomando $\delta > 0$ lo suficientemente pequeño, tenemos que el $\frac{\lambda^+}{2} - \frac{|o(t^2)|}{t^2}$ será estrictamente positivo para todo $t \in (0, \delta)$, y por lo tanto f no puede ser un máximo local. Análogamente,

$$\begin{aligned}
f(x + tv^-) - f(x) &= +tDf(x)v^- + \frac{1}{2}\langle tv^-, D^2f(x)tv^-\rangle + o(t^2) \\
&= \frac{t^2}{2}\lambda^- o(t^2) \\
&= t^2\left(\frac{\lambda^-}{2} + \frac{|o(t^2)|}{t^2}\right).
\end{aligned}$$

Tomando otro valor $\delta > 0$ lo suficientemente pequeño, tenemos que el $\frac{\lambda^-}{2} + \frac{|o(t^2)|}{t^2}$ será estrictamente negativo para todo $t \in (0, \delta)$, y por lo tanto f no puede ser un mínimo local. Concluimos entonces que x es punto silla, lo que termina la demostración. □

Atención

Para un punto crítico x de f, tal que $D^2 f(x)$ sea solo semidefinida positiva, el teorema anterior no nos permite concluir nada. Es decir, el punto en cuestión podría ser un mínimo local o un punto silla, y no tendríamos como decidir. Lo mismo pasa si $D^2 f(x)$ sea solo semidefinida negativa.

3.4 Funciones convexas

Para una función $f : \mathbb{R}^n \to \mathbb{R}$ y para un punto $x^* \in \mathbb{R}^n$, sabemos que

$$x^* \text{ es mínimo global de } f \implies x^* \text{ es mínimo local de } f \implies x^* \text{ es punto crítico de } f. \tag{3.22}$$

Para la mayoría de las funciones, las reciprocas de estas implicancias no son ciertas. Es decir, x^* puede ser un punto crítico de f, sin ser un mínimo local de f, y a la vez podría ser un mínimo local de f sin ser un mínimo global. Sin embargo, hay una familia de funciones donde estos tres tipos de puntos coinciden: las funciones convexas.

Definición 3.4.1 (Función convexa)**.** *Una función $f : \mathbb{R}^n \to \mathbb{R}$ se dice **convexa** si para todo $x, y \in \mathbb{R}^n$, y para todo $\lambda \in [0, 1]$ se cumple que*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y). \tag{3.23}$$

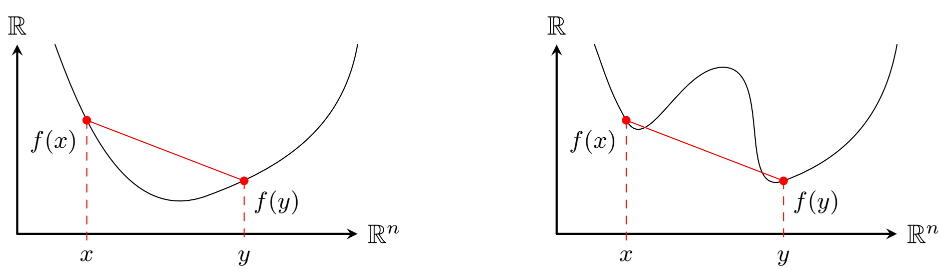


Figura 3.7: El segmento rojo corresponde a $[(x, f(x)), (y, f(y))]$ en $\mathbb{R}^{n+1}$, o equivalentemente, al grafo de la función $\lambda \in [0, 1] \mapsto \lambda f(x) + (1 - \lambda) f(y)$. A la izquierda, la función es convexa. A la derecha, la función no es convexa.

La propiedad de convexidad tiene una interpretación geométrica. Para todo par de puntos $x, y \in \mathbb{R}^n$, el segmento en $\mathbb{R}^{n+1}$ que une los puntos $(x, f(x))$ y $(y, f(y))$ queda por sobre el grafo de la función f. Esto se ilustra en la Figura 3.7. Esto nos dice que de hecho la convexidad es una propiedad *geométrica* de la función.

Proposición 3.4.2. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa y $x \in \mathbb{R}^n$. Entonces se tiene que*

$$x \textit{ es mínimo local de } f \implies x \textit{ es mínimo global de } f$$

Demostración. Sea x un mínimo local de f. Tomemos $\delta > 0$ tal que $f(x) \leq f(y)$ para todo $y \in B(x, \delta)$. Tomemos $z \in \mathbb{R}^n$ y tomemos $\lambda \in (0, 1)$ lo suficientemente pequeño tal que

$$y = (1 - \lambda)x + \lambda z \in B(x, \delta).$$

Usando la convexidad de f y el hecho que x es mínimo local, podemos escribir que

$$f(x) \leq f(y) = f((1-\lambda)x + \lambda z) \leq (1-\lambda)f(x) + \lambda f(z).$$

Luego, restando $(1-\lambda)f(x)$ a amobs lados de la desigualdad, y usando que $\lambda > 0$, tenemos que

$$\lambda f(x) \leq \lambda f(z) \overset{\lambda>0}{\implies} f(x) \leq f(z).$$

Como z es un punto arbitrario, concluimos que

$$\forall z \in \mathbb{R}^n, f(x) \leq f(z),$$

es decir, x es un mínimo global de f. Esto concluye la demostración. □

Para funciones diferenciables, podemos caracterizar la convexidad como una propiedad analítica del gradiente.

Proposición 3.4.3 (Caracterización de primer orden)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función diferenciable. las siguientes afirmaciones son equivalentes:*

(i) f *es convexa.*

(ii) Para todo $x, y \in \mathbb{R}^n$, $f(x) + \langle \nabla f(x), y - x\rangle \leq f(y)$.

(iii) Para todo $x, y \in \mathbb{R}^n$, $\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq 0$.

Demostración. Demostraremos el ciclo de implicancias.

$(i) \implies (ii)$**:** Tomemos $x, y \in \mathbb{R}^n$ distintos entre si, y fijemos $d = y - x$. Consideremos la sucesión $\lambda_k = 1/k$. Tenemos que claramente $\lambda_k \leq 1$ para todo $k \in \mathbb{N}$ y además $\lambda_k \to 0$. Luego, como f es diferenciable, podemos escribir:

$$\begin{aligned}
f(x) + \langle \nabla f(x), y - x\rangle &= \lim_{\lambda\to 0} f(x) + \frac{f(x+\lambda d) - f(x)}{\lambda} \\
&= \lim_{\lambda\to 0} f(x) + \frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda} \\
&= \lim_{k\to\infty} f(x) + \frac{f((1-\lambda_k)x + \lambda_k y) - f(x)}{\lambda_k} \\
&\leq \lim_{k\to\infty} f(x) + \frac{(1-\lambda_k)f(x) + \lambda_k f(y) - f(x)}{\lambda_k} \\
&= \lim_{k\to\infty} f(x) + \frac{\lambda_k f(y) - \lambda_k f(x)}{\lambda_k} = f(y).
\end{aligned}$$

Como x e y son arbitrarios, se concluye (ii).

$(ii) \implies (iii)$**:** Sean $x, y \in \mathbb{R}^n$ arbitrarios. Usando (ii) tenemos que

$$\begin{aligned}
\langle \nabla f(x), y - x\rangle &\leq f(y) - f(x) \\
\langle \nabla f(y), x - y\rangle &\leq f(x) - f(y)
\end{aligned}$$

Sumando ambas desigualdades tenemos que

$$\langle \nabla f(x), y - x\rangle + \langle \nabla f(y), x - y\rangle \leq 0.$$

Luego, podemos escribir

$$\begin{aligned}
0 &\leq -\langle \nabla f(x), y - x\rangle - \langle \nabla f(y), x - y\rangle \\
&= \langle \nabla f(x), x - y\rangle + \langle -\nabla f(y), x - y\rangle \\
&= \langle \nabla f(x) - \nabla f(y), x - y\rangle.
\end{aligned}$$

Como $x, y \in \mathbb{R}$ son arbitrarios, se concluye (iii).

(iii) $\implies$ *(i)*: Tomemos $x, y \in \mathbb{R}^n$ y $\lambda \in [0,1]$. Queremos demostrar que la desigualdad (3.23). Para esto, definamos $z = \lambda x + (1-\lambda)y$ y, razonando por contradicción, supongamos que $f(z) > \lambda f(x) + (1-\lambda)f(y)$.
Consideremos la función $g(t) = f(y + t(x-y))$ y la función $\ell(t) = f(y) + t(f(x) - f(y))$. Notemos que $g'(t) = \langle \nabla f(y + t(x-y)), x - y\rangle$, para todo $t \in [0,1]$. Además, tenemos que $g(0) = \ell(0)$, $g(1) = \ell(1)$, y que $g(\lambda) > \ell(\lambda)$. Usando el teorema del valor medio, podemos escribir lo siguiente:

1. Existe $\lambda_1 \in (0, \lambda)$ tal que
$$g'(\lambda_1) = \frac{g(\lambda) - g(0)}{\lambda} > \frac{\ell(\lambda) - \ell(0)}{\lambda} = f(x) - f(y).$$

2. Existe $\lambda_1 \in (0, \lambda)$ tal que
$$g'(\lambda_2) = \frac{g(1) - g(\lambda)}{1-\lambda} < \frac{\ell(1) - \ell(\lambda)}{1-\lambda} = f(x) - f(y).$$

Definamos ahora $z_1 = \lambda_1 x + (1-\lambda_1)y$ y $z_2 = \lambda_2 x + (1-\lambda_2)y$. Primero, notemos que
$$z_2 - z_1 = (\lambda_2 - \lambda_1)(x-y).$$
Con esto en consideración, notando que $\lambda_2 - \lambda_1 > 0$, podemos escribir
$$\begin{aligned}
0 > g'(\lambda_2) - g'(\lambda_1) &= \langle \nabla f(z_2), x-y\rangle - \langle \nabla f(z_1), x-y\rangle \\
&= \langle \nabla f(z_2) - \nabla f(z_1), x-y\rangle \\
&= \frac{1}{\lambda_2 - \lambda_1}\langle \nabla f(z_2) - \nabla f(z_1), (\lambda_2 - \lambda_1)(x-y)\rangle = \frac{1}{\lambda_2 - \lambda_1}\langle \nabla f(z_2) - \nabla f(z_1), z_2 - z_1\rangle.
\end{aligned}$$
Multiplicando por $\lambda_2 - \lambda_1$, concluimos entonces que
$$\langle \nabla f(z_2) - \nabla f(z_1), z_2 - z_1\rangle < 0,$$
lo cual es una contradicción con *(iii)*. Por lo tanto, tenemos que $f(z) \leq \lambda f(x) + (1-\lambda)f(y)$, que es lo que queríamos demostrar. Por lo tanto, concluimos *(i)*, lo que concluye la demostración. □

Para funciones de clase $\mathcal{C}^2$, también podemos caracterizar la convexidad en términos de la matriz Hessiana.

Proposición 3.4.4 (Caracterización de segundo orden)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función de clase $\mathcal{C}^2$. Se tiene que*
$$D^2 f(x) \text{ es semidefinida positiva, para todo } x \in \mathbb{R}^n \iff f \text{ es convexa.}$$

Demostración. Estudiemos la doble implicancia:

$\implies$**:** Sean $x, z \in \mathbb{R}^n$. Ocupando el teorema de expansión de Taylor de segundo orden (ver Teorema 2.5.7, sabemos que
$$f(z) = f(x) + \langle \nabla f(x), z - x\rangle + \frac{1}{2}\langle z - x, D^2 f(x_t)(z-x)\rangle,$$
con $x_t = tz + (1-t)x$, para algún $t \in (0,1)$. Luego, como $D^2 f(x_t)$ es semidefinida positiva, se tiene que
$$f(z) \geq f(x) + \langle \nabla f(x), z - x\rangle,$$
lo que muestra la convexidad de f aplicando la Proposición 3.4.3.

$\Longleftarrow$ **:** Razonemos por contradicción, y supongamos que existe $x, z \in \mathbb{R}^n$ tal que $\langle z, D^2 f(x) z\rangle < 0$. Como f es de clase $\mathcal{C}^2$, la matriz Hessiana es continua, y por lo tanto existe $\delta > 0$ tal que para todo $t \in [0, \delta]$, se tiene que $\langle z, D^2 f(x + tz) z\rangle < 0$. Tomando $h = \delta z$, tenemos que

$$\forall t \in [0,1], \langle h, D^2 f(x + th) h\rangle = \delta^2 \langle z, D^2 f(x + (\delta t) z) z\rangle < 0.$$

Luego, ocupando nuevamente el Teorema de Taylor, exite $t \in (0,1)$ tal que

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\langle h, D^2 f(x+th) h\rangle < f(x) + \langle \nabla f(x), h\rangle.$$

Esto último es una contradicción con que f sea convexa, en vista de la Proposición 3.4.3. □

Para terminar, usando la caracterización de primer orden de la convexidad, podemos mostrar el resultado principal de esta sección: todo punto crítico de una función convexa es un mínimo global.

Teorema 3.4.5. *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa y $x \in \mathbb{R}^n$. Se tiene que*

$$x \text{ es punto crítico de } f \implies x \text{ es mínimo global de } f.$$

Demostración. Como f es convexa, la Proposición 3.4.3 nos asegura que

$$\forall y \in \mathbb{R}^n, f(x) + \langle \nabla f(x), y - x\rangle \leq f(y).$$

Como x es punto crítico de f, tenemos que $\nabla f(x) = 0$, y por lo tanto, concluimos que

$$\forall y \in \mathbb{R}^n, f(x) \leq f(y).$$

Es decir, x es mínimo global de f. □

3.5 Optimización con restricciones

Hasta ahora, nuestra única herramienta para resolver problemas de optimización es la Regla de Fermat (ver Teorema 3.3.1): Si el problema no tiene restricciones, resolvemos el sistema $\nabla f(x) = 0$ para identificar los candidatos a mínimos y máximos (locales y globales). Sin embargo, en la práctica, los problemas de optimización tienen restricciones y por lo tanto la Regla de Fermat no es válida.

Ejemplo 3.5.1. Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x,y) = x^2 + y^2$, y consideremos el problema

$$\begin{cases} \max\limits_{x,y} & f(x,y) \\ s.a. & x^2 + y^2 = 1. \end{cases}$$

En este caso, todos los puntos factibles son máximos y mínimos globales, pues f es constante en la región factible $X = \{(x,y) : x^2 + y^2 = 1\}$. Sin embargo $\nabla f(x,y) = (2x, 2y)$, y por lo tanto $\nabla f(x,y) = 0$ solo se verifica en $(x,y)0(0,0)$ que no es un punto factible. ◊

En este curso estudiaremos la técnica de *multiplicadores de Lagrange*, que se enfoca en problemas únicamente con restricciones de igualdad. Es decir, para $f : \mathbb{R}^n \to \mathbb{R}$ y $h : \mathbb{R}^n \to \mathbb{R}^m$ diferenciables, consideraremos un problema de la forma

$$(P) = \begin{cases} \min_{x \in \mathbb{R}^n} & f(x) \\ s.a. & \begin{cases} h_1(x) = 0, \\ \vdots \\ h_m(x) = 0. \end{cases} \end{cases} \tag{3.24}$$

Vamos también a suponer que $n > m$, es decir, que tenemos más variables que restricciones.

Nota: Problemas con restricciones de desigualdad

También es posible trabajar con restricciones de desigualdad, utilizando lo que se conoce como el *Teorema de Karush-Kuhn-Tucker*. Sin embargo, esto está fuera de los alcances del curso.

La idea princpipal de la técnica de multiplicadores de Lagrange es definir una función auxiliar, llamada el Lagrangiano del problema, y aplicar la regla de Fermat a esta función.

Definición 3.5.2 (Lagrangiano)**.** *Consideremos el problema* (3.24)*. Se define el Lagrangiano de este problema como*

$$\begin{aligned}\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m &\to \mathbb{R}\\ (x, \lambda) &\mapsto f(x) - \sum_{j=1}^{m} \lambda_j h_j(x).\end{aligned}$$

Las variables auxiliares $\lambda \in \mathbb{R}^m$ reciben el nombre de multiplicadores de Lagrange.

La regla de Fermat aplicada al Lagrangiano consiste en resolver el sistema $\nabla\mathcal{L}(x, \lambda) = 0$. Notemos que

$$\nabla\mathcal{L}(x, \lambda) = \begin{pmatrix}\nabla_x\mathcal{L}(x, \lambda)\\ \nabla_\lambda\mathcal{L}(x, \lambda)\end{pmatrix} = \begin{pmatrix}\nabla f(x) - \sum_{j=1}^{m} \lambda_j \nabla h_j(x)\\ -h(x)\end{pmatrix}$$

Por lo tanto, resolver $\nabla\mathcal{L}(x, \lambda) = 0$ equivale a resolver el sistema

$$\begin{aligned}\nabla f(x) - \textstyle\sum_{j=1}^{m} \lambda_j \nabla h_j(x) &= 0,\\ h(x) &= 0.\end{aligned} \tag{3.25}$$

El siguiente teorema nos dice cuando podemos usar el sistema (3.25) para buscar soluciones del problema (3.24). Enunciaremos el teorema sin demostración, debido a que no contamos con las herramientas teóricas necesarias para demostrarlo.

Teorema 3.5.3 (Teorema de multiplicadores de Lagrange)**.** *Consideremos el problema*

$$(P) = \begin{cases}\min_{x\in\mathbb{R}^n} & f(x)\\ s.a. & \begin{cases}h_1(x) = 0,\\ \quad\vdots\\ h_m(x) = 0.\end{cases}\end{cases}$$

y supongamos que x^* es un óptimo local (mínimo o máximo). Si además el conjunto de vectores $\{\nabla h_1(x^*), \ldots, \nabla h_m(x^*)\}$ es ***linealmente independientes****, entonces existe $\lambda^* \in \mathbb{R}^m$ tal que el sistema* (3.25) *se verifica en (x^*, λ^*). Es decir, tal que*

$$\begin{aligned}\nabla f(x^*) - \textstyle\sum_{j=1}^{m} \lambda_j^* \nabla h_j(x^*) &= 0,\\ h(x^*) &= 0.\end{aligned}$$

El teorema anterior nos dice que si el problema (3.24) tiene solución x^*, entonces esa solución aparecerá como punto crítico del Lagrangiano $\mathcal{L}$ (con algún multiplicador λ^*), siempre y cuando se verifique la condición de calificación de que los gradientes $\{\nabla h_1(x^*), \ldots, \nabla h_m(x^*)\}$ sean linealmente independientes. Entonces, para ocupar el teorema de multiplicadores de Lagrange, debemos seguir los siguientes pasos:

1. Verifico que el problema (3.24) tenga solución (usando tipicamente el Teorema de Weiestrass, ver Teorema 3.2.6).
2. Verifico que para todo $x \in \mathbb{R}^n$ que satisface las restricciones $h(x) = 0$, el conjunto de vectores $\{\nabla h_1(x), \ldots, \nabla h_m(x)\}$ es linealmente independiente. Si algún punto falla esta condición, **lo guardo**.
3. Resuelvo el sistema $\nabla \mathcal{L}(x, \lambda) = 0$, que corresponde al sistema (3.25). **Guardo todos los puntos que resuelven el sistema**.
4. El Teorema 3.5.3 garantiza el mínimo se debe encontrar entre los puntos guardados: Los evalúo en la función objetivo f y me quedo con el que alcance el menor valor.

Atención

El sistema $\nabla \mathcal{L}(x, \lambda) = 0$ encuentra los puntos críticos del problema (3.24), que son candidatos a solución. Sin embargo, si un punto x no verifica que $\{\nabla h_1(x), \ldots, \nabla h_m(x)\}$ es linealmente independiente, podría ser solución del problema de optimización, pero no aparecer como punto crítico. Por eso en el paso 2 guardamos como posibles candidatos los puntos que no verifican la condición de calificación de independencia lineal.

3.6 Selección de Problemas

1. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ dada por
$$f(x, y) = xy(1 - x - y).$$
Determine los puntos críticos de f y clasifíquelos según sean máximos, mínimos o puntos silla. Justifique adecuadamente sus respuestas.
2. Verifique que el punto $(x_0, y_0, z_0) = (1, 1, 1)$ es un punto crítico de la función
$$g(x, y, z) = x^4 + y^4 + z^4 - 4xyz$$
y determine si se trata de un máximo, un mínimo o un punto silla.
3. Considere la función dada por
$$f(x, y) = x^2 - \operatorname{sen}(x + y).$$
   a) Encuentre todos sus puntos críticos.
   b) Determine la naturaleza de estos puntos (máximos o mínimos locales, puntos silla).
4. En esta pregunta, deseamos encontrar los valores máximos y mínimos de la función
$$f(x, y) = 4x^2 + 10y^2,$$
sobre el conjunto de restricciones $D = \{(x, y) \in \mathbb{R}^2 \ : \ x^2 + y^2 \leq 4\}$. Para poder llevar esto a cabo, se le propone la siguiente estrategia de resolución:
   a) Optimice f sin restricciones, y determine si el(los) punto(s) que encuentra pertenece(n) o no al interior de D. Determine además la naturaleza de estos puntos.
   b) Usando Multiplicadores de Lagrange, encuentre los máximos y mínimos de f en la frontera de D, donde se cumple que $x^2 + y^2 = 4$.

c) Compare todos los puntos encontrados y concluya.

5. La conveniencia social de una empresa, con frecuencia incluye elegir entre la ventaja comercial de la empresa y las desventajas sociales y ecológicas que pueda generar. Por ejemplo, la industria de la madera proporciona productos de papel a la sociedad e ingresos a muchos trabajadores y empresas, pero los beneficios pueden contrarrestarse con la destrucción del hábitat de especies en vías de extinción. Suponga que la conveniencia social de una empresa particular, se mide mediante la función

$$D(x,y) = (16-6x)x - (y^2 - 4xy + 40), \quad x \geq 0,\ y \geq 0,$$

donde x representa las ventajas comerciales e y el impacto ecológico. La empresa se considera *deseable* si $D \geq 0$ e *indeseable* si $D < 0$. ¿Qué valores de x e y maximizarán la conveniencia social? Interprete sus resultados, ¿es posible que esta empresa sea *deseable* socialmente?

6. Suponga que usted es el fabricante de un nuevo celular de gama media, y desea venderlo a \$350.000 la unidad. Luego de hacer un estudio de mercado, usted estima que si gasta x millones de pesos en desarrollo e y millones de pesos en publicidad, los consumidores comprarán

$$u(x,y) = \frac{250y}{y+2} + \frac{100x}{x+5}$$

unidades del producto, aproximadamente. Si cada celular cuesta \$150.000 en producirse, ¿cuánto dinero destinaría usted para desarrollo y para publicidad, si desea maximizar sus ganancias totales?

7. Una lechería produce leche entera y leche descremada en cantidades x e y galones, respectivamente. Suponga que el precio de la leche entera es

$$p(x) = 100 - x$$

y el de la leche descremada es

$$q(y) = 100 - y.$$

Suponga además que

$$C(x,y) = x^2 + xy + y^2$$

es la función de costos conjuntos de los artículos. ¿Cuáles deberían ser los valores de x e y para maximizar las utilidades?

8. Una rara enfermedad puede ser tratada administrando 70 [ml] de un medicamento C, pero provocando efectos colaterales graves en las personas en dosis de ese nivel o superiores. En busca de un método más seguro para tratar a sus pacientes, un médico usa dos medicamentos combinados, A y B, que no tienen efectos colaterales mientras su dosis combinada **sea menor a 60 [ml]**.

   Luego de muchos estudios, el médico descubre que al administrar x [ml] de A e y [ml] de B, el tratamiento es equivalente a E [ml] del medicamento C, donde

$$E(x,y) = 0.05(xy - x(2x-95) + y(20-y)),$$

   con la garantía de que no genera efectos colaterales, aunque $E \geq 70$; siempre y cuando la dosis combinada $x + y \leq 60$.

   a) Determine el(los) punto(s) crítico(s) de $E(x,y)$.

   b) A partir del (de los) punto(s) encontrados en la parte anterior, ¿qué dosis de A y B **maximizan** el nivel equivalente E del medicamento C?

c) Si el médico administra las dosis óptimas encontradas en la parte anterior, el efecto combinado, ¿es suficiente para ayudar al paciente sin correr el riesgo de sufrir efectos colaterales?

9. Al editor de un libro se le han asignado \$60.000.000 para invertir en el desarrollo y la promoción de un nuevo libro. Se calcula que si se gastan x millones de pesos en desarrollo e y millones de pesos en publicidad, se venderán

$$g(x, y) = 20x^{3/2}y$$

ejemplares del libro en cuestión. ¿Cuánto dinero se debe asignar a cada ítem para maximizar las ventas?

10. El área de descanso para automovilistas que planea construir una autopista al lado de una de sus carreteras principales, debe ser rectangular y tener un área de 5000 metros cuadrados, cercada por los tres lados no adyacentes a la carretera, como muestra la figura adjunta.

Figura 3.8: Área de Descanso

¿Cuál es la menor cantidad de cercado necesaria para completar el trabajo?

Capítulo 4

Integración en $\mathbb{R}^n$

Consideremos una función $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$. Cuando $n = 1$ y $D = [a, b]$, recordemos que la integral definida

$$\int_a^b f(x)dx$$

representa, cuando $f(x) \geq 0$, el *área bajo la curva del grafo de* f, esto es, el área del conjunto

$$S = \{(x, y) \in \mathbb{R}^2 \ : \ x \in [a, b], \ 0 \leq y \leq f(x)\}.$$

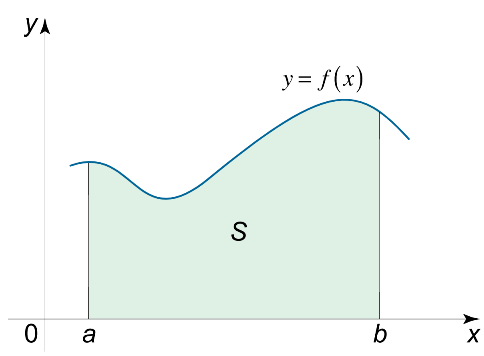


Figura 4.1: Representación gráfica de S, área bajo la curva de $f : [a, b] \to \mathbb{R}$.

A lo largo de las siguientes páginas, trataremos de extender esta noción de la forma más natural posible, al contexto de varias variables. Si pensamos en $n = 2$, recordemos que el grafo de una función puede ser interpretado como una especie de *superficie* en $\mathbb{R}^3$ (concepto que definiremos más cabalmente en el Capítulo 5) y, en consecuencia, sería razonable definir la cantidad

$$\iint_D f(x, y) \ dxdy$$

como el *volumen bajo el grafo de* f, siempre que $f(x, y) \geq 0$ para $x \in D$. Vale decir, el volumen del cuerpo geométrico

$$V = \{(x, y, z) \in \mathbb{R}^3 \ : \ (x, y) \in D, \ 0 \leq z \leq f(x, y)\}.$$

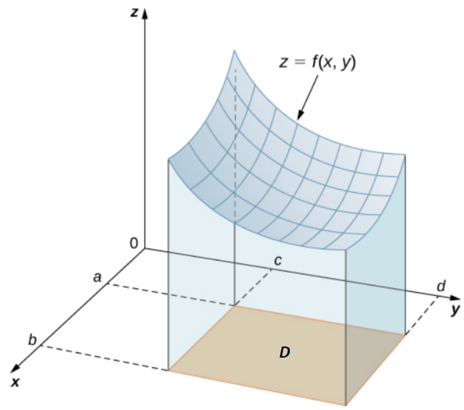


Figura 4.2: Representación gráfica de V, volumen bajo el grafo de $f : D \to \mathbb{R}$.

4.1 Definiciones Fundamentales

Comencemos estudiando un tipo sencillo de conjuntos, que nos permitirán sentar las bases del concepto de integral a desarrollar en esta sección.

Definición 4.1.1 (Rectángulo)**.** *Definimos un* rectángulo de lados paralelos a los ejes coordenados *como cualquier conjunto $R \subseteq \mathbb{R}^n$ de la forma*

$$\begin{aligned} R &= \{x \in \mathbb{R}^n \ : \ a_i \leq x_i \leq b_i \quad \textit{para todo } i = 1, \dots, n\} \\ &= [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n] = \prod_{i=1}^{n} [a_i, b_i]. \end{aligned}$$

Definimos adicionalmente el área de un rectángulo R *como*

$$A(R) = \prod_{i=1}^{n} (b_i - a_i).$$

Atención

Aunque un rectángulo, bajo la definición anterior, podría tener tres o más dimensiones (y por ende corresponder a un paralelepípedo, por ejemplo) dado que es un subconjunto de $\mathbb{R}^n$; usaremos en cualquier contexto el concepto de área para referirnos a su unidad de medida.

Sea $R \subseteq \mathbb{R}^n$ un rectángulo. Definiremos ahora un concepto similar al de partición en el caso de la integral en $\mathbb{R}$, que nos permitirá a su vez definir en este nuevo contexto las sumas superiores e inferiores, de manera análoga a lo que ya conocemos en la recta real.

Definición 4.1.2 (Reticulado)**.** *Un reticulado $\mathcal{S}$ de un rectángulo $R \subseteq \mathbb{R}^n$ es una familia de conjuntos*

$$\mathcal{S} = \{R_i\}_{i \in I} \subseteq \mathcal{P}(\mathbb{R}^n),$$

donde I es un conjunto de índices, y cada $R_i \in \mathcal{S}$ es a su vez un rectángulo; tal que cumple las siguientes condiciones:

1. $int(R_i) \cap int(R_j) = \varnothing, \quad \forall i \neq j.$

2. $R = \bigcup_{i \in I} R_i$.

Ejemplo 4.1.3. Si consideramos $R = [a,b] \times [c,d] \in \mathbb{R}^2$ y el conjunto de índices

$$I = \{(i,j) \ : \ i,j \in \mathbb{N}, \ 1 \leq i \leq 8, \ 1 \leq j \leq 9\},$$

la figura adjunta representa un reticulado de R. ◊

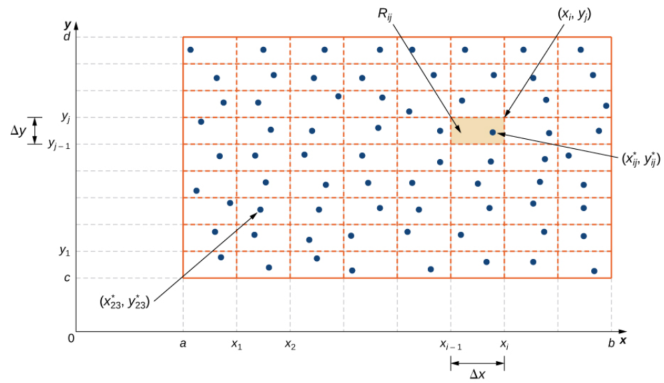


Figura 4.3: Reticulado Uniforme para $[a,b] \times [c,d]$.

Definición 4.1.4 (Sumas de Riemann)**.** *Sea $R \subseteq \mathbb{R}^n$ un rectángulo y un reticulado del mismo, $\mathcal{S}$. Consideremos además $f : R \to \mathbb{R}$ una función* acotada*.*

- *Definimos la* suma inferior *de f sobre el reticulado $\mathcal{S}$, como*

$$\mathcal{I}_S(f) = \sum_{i \in I} m_{R_i}(f) \cdot A(R_i),$$

donde $m_{R_i}(f) = \inf_{x \in \mathbb{R}_i} f(x)$.

- *Definimos la* suma superior *de f sobre el reticulado $\mathcal{S}$, como*

$$\mathcal{S}_S(f) = \sum_{i \in I} M_{R_i}(f) \cdot A(R_i),$$

donde $M_{R_i}(f) = \sup_{x \in \mathbb{R}_i} f(x)$.

Esta definición de sumas superiores e inferiores, permite hacer cálculos similares a los que ya conocemos para una variable. Notemos que en el caso de $n = 2$, el cálculo

$$m_{R_i}(f) \cdot A(R_i)$$

representa el *volumen* de un paralelepípedo de base R_i y altura $m_{R_i}(f)$, y al sumar en todos los rectángulos R_i del reticulado, estamos obteniendo una aproximación inferior del volumen debajo de la superficie generada por el grafo de f, delimitado por R. Por otra parte, el cálculo

$$M_{R_i}(f) \cdot A(R_i)$$

corresponde a algo similar, pero considerando altura $M_{R_i}(f)$ para los cálculos, obteniendo así una aproximación superior sumando en todo el reticulado.

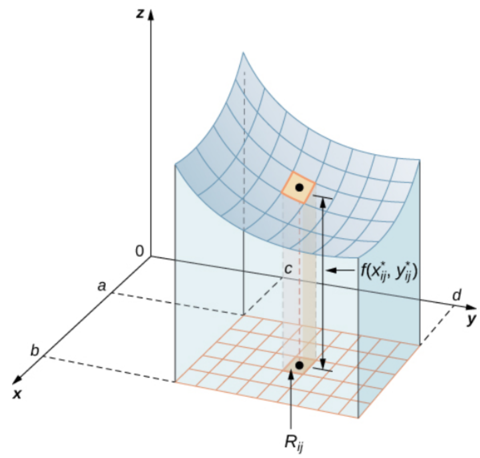


Figura 4.4: Aproximación del cálculo del volumen bajo $z = f(x,y)$ y delimitado por $[a,b] \times [c,d]$.

Definición 4.1.5. *Si tenemos dos reticulados $\mathcal{S}_1$ y $\mathcal{S}_2$ de un mismo rectángulo R, diremos que $\mathcal{S}_2$ es más fino que $\mathcal{S}_1$, si todos los rectángulos de $\mathcal{S}_2$ están contenidos en algún triángulo de $\mathcal{S}_1$.*

No es muy difícil convencerse de que, a medida que vamos *refinando* los rectángulos del reticulado $\mathcal{S}$, las sumas inferior y superior aproximan cada vez de mejor forma el volumen que buscamos calcular.

Definición 4.1.6 (Integrabilidad)**.** *Si existe una sucesión de reticulados $\{S_N\}_{N\in\mathbb{N}}$ tal que*

$$\lim_{N\to\infty} \mathcal{S}_{S_N}(f) - \mathcal{I}_{S_N}(f) = 0,$$

*diremos que f es una función **integrable** en R, y además*

$$\int_R f(x)dx = \lim_{N\to\infty} \mathcal{S}_{S_N}(f) = \lim_{N\to\infty} \mathcal{I}_{S_N}(f).$$

Ejemplo 4.1.7. Calcular sumas superiores e inferiores, incluso en el caso de funciones sencillas de trabajar, puede ser un poco tedioso. Desarrollamos acá un ejemplo ilustrativo, para la función

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

en el rectángulo $R = [0,1] \times [0,1]$. Consideremos para $N \in \mathbb{N}$ el reticulado $\mathcal{S}_N$ dado por los rectángulos

$$R_{ij}^N = \left[\frac{i}{N}, \frac{i+1}{N}\right] \times \left[\frac{j}{N}, \frac{j+1}{N}\right],$$

para $0 \leq i,j \leq N-1$, cuya área está dada por

$$A(R_{ij}^N) = \frac{1}{N^2}.$$

En este caso,

$$\mathcal{S}_{S_N}(f) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} A(R_{ij}^N) \cdot M_{R_{ij}^N}(f),$$

y por la función elegida, podemos notar que

$$M_{R_{ij}^N}(f) = \frac{i+1}{N} + \frac{j+1}{N}.$$

En consecuencia,

$$\begin{aligned}
\mathcal{S}_{S_N}(f) &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\frac{i+j+2}{N^3}\\
&= \frac{1}{N^3}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}(i+j+2)\\
&= \frac{1}{N^3}\sum_{i=0}^{N-1}\left(iN+\frac{(N-1)N}{2}+2N\right)\\
&= \frac{1}{N^3}\left(\frac{N^2(N-1)}{2}+\frac{(N-1)N^2}{2}+2N^2\right)\\
&= \frac{N^2(N+1)}{N^3} = 1+\frac{1}{N}.
\end{aligned}$$

De manera similar, notando que en el caso de la suma inferior se puede obtener

$$m_{R_{ij}^N}(f) = \frac{i}{N}+\frac{j}{N},$$

al desarrollar, obtenemos

$$\mathcal{I}_{S_N}(f) = \frac{N^2(N-1)}{N^3} = 1-\frac{1}{N}.$$

De esta manera, cuando $N\to\infty$,

$$\lim_{N\to\infty}\mathcal{S}_{S_N}-\mathcal{I}_{S_N} = \lim_{N\to\infty}\frac{2}{N} = 0,$$

por lo que f es, en efecto, integrable en $[0,1]\times[0,1]$. Más aún, hemos probado que

$$\iint\limits_{[0,1]\times[0,1]} f(x,y)\;dxdy = \lim_{N\to\infty}\mathcal{S}_{S_n}(f) = 1.$$

$\Diamond$

Enunciamos a continuación un primer resultado importante, que nos permitirá estar seguros de la integrabilidad de una función f bajo ciertas condiciones elementales, sin la necesidad de pasar por sus Sumas de Riemann para verificarlo.

Teorema 4.1.8. *Sea $f : R\subseteq\mathbb{R}^n\to\mathbb{R}$ una función continua, donde R es un rectángulo. Entonces, f es integrable.*

Demostración. Necesitamos el siguiente resultado intermedio:

Proposición Auxiliar. Sea $f : K\subset\mathbb{R}^N\to\mathbb{R}$ una función continua, donde K es un conjunto cerrado y acotado. Entonces f es uniformemente continua, es decir:

$$\forall\varepsilon>0,\ \exists\delta>0 \text{ tal que } \|x-y\|<\delta\Rightarrow|f(x)-f(y)|<\varepsilon\quad\forall x,y\in K.$$

Demostración de la Proposición Auxiliar. Supongamos que esta propiedad no es válida. Entonces existe $\varepsilon_0>0$ tal que para todo $n\in\mathbb{N}$ existen sucesiones $\{x_n\},\{y_n\}\subset K$ tales que:

$$\|x_n-y_n\|<\frac{1}{n},\quad\text{pero}\quad|f(x_n)-f(y_n)|\geq\varepsilon_0.$$

Como $\{x_n\} \subset K$, que es cerrado y acotado, tiene una subsucesión convergente $x_{n_j} \to \bar{x} \in K$. Entonces también $y_{n_j} \to \bar{x}$. Como f es continua:

$$\lim_{j\to\infty} f(x_{n_j}) = \lim_{j\to\infty} f(y_{n_j}) = f(\bar{x}),$$

lo cual contradice que $|f(x_{n_j}) - f(y_{n_j})| \geq \varepsilon_0$. ■

Demostración del Teorema: En virtud de la Proposición Auxiliar, basta demostrar que existe una sucesión de reticulados S_n tal que:

$$S_{S_n}(f) - I_{S_n}(f) \to 0.$$

Consideremos, para $n \geq 1$, el reticulado uniforme dado por:

$$\left[a_1 + \frac{k_1 - 1}{n}(b_1 - a_1),\, a_1 + \frac{k_1}{n}(b_1 - a_1)\right] \times \cdots \times \left[a_N + \frac{k_N - 1}{n}(b_N - a_N),\, a_N + \frac{k_N}{n}(b_N - a_N)\right]$$

con $0 \leq k_j \leq n$ para todo $j = 1, \ldots, N$. Sea R_i la enumeración de estos rectángulos para $i = 1, \ldots, n^N$. Entonces, si $x, y \in R_i$, se tiene:

$$\|x - y\| < \frac{C}{\sqrt{n}}.$$

Por continuidad uniforme de f, para cada $m > 1$ existe $n = n_m$ suficientemente grande tal que:

$$|f(x) - f(y)| < \frac{1}{m} \quad \forall x, y \in R_i.$$

En particular:

$$M_{R_i}(f) \leq m_{R_i}(f) + \frac{1}{m}, \quad \Rightarrow \quad S_{S_m} \leq I_{S_m} + \frac{1}{m} \cdot V(R).$$

Como m es arbitrario, se concluye que f es integrable. □

Si bien el resultado anterior es de gran importancia, no es suficiente en la práctica para el cálculo de integrales; dado que por ejemplo, nos interesará definir la integral de funciones que no sean continuas, o hacerlo en conjuntos que no sean rectángulos. Comencemos a mirar esta segunda situación.
Sea $D \subseteq \mathbb{R}^n$ un conjunto acotado, no necesariamente rectangular, y una función acotada $f : D \to \mathbb{R}$. Deseamos definir la cantidad

$$\int_D f(x)\, dx$$

como aquella correspondiente al volumen de la región entre la base D y el gráfico de f. Para lograr esto, basta considerar la función f_D dada por

$$f_D(x) = \begin{cases} f(x) & \text{si } x \in D \\ 0 & \text{si } x \notin D \end{cases},$$

y R un rectángulo tal que $D \subseteq R$.

Definición 4.1.9 (Integrabilidad sobre un dominio no rectangular)**.** *Diremos que f es integrable sobre el conjunto D, si f_D es integrable sobre R. En este caso, escribiremos*

$$\int_D f(x)dx = \int_R f_D(x)dx.$$

Un caso especial de lo anterior será el de la función constante $f(x) \equiv 1$. En este caso, f_D no será continua, pero normalmente sí será integrable.

Definición 4.1.10 (Volumen de una región D)**.** *Cuando la cantidad*

$$V(D) = \int_D 1 \; dx$$

esté bien definida, diremos que esta representa el volumen *de la región de integración D.*

- *Cuando $n = 1$, interpretaremos esto como el* largo *de D.*
- *Cuando $n = 2$, interpretaremos esto como el* área *de D.*

Ya estamos en condiciones de enunciar las propiedades elementales de la integral.

Proposición 4.1.11 (Propiedades de la integral)**.** *Sean f y g dos funciones integrables en una misma región $D \subseteq \mathbb{R}^n$. Entonces,*

- *Para $\lambda \in \mathbb{R}$, la función $\lambda f + g$ también es integrable sobre D, y*
$$\int_D (\lambda f(x) + g(x))dx = \lambda \int_D f(x)dx + \int_D g(x)dx.$$
- *Si $f(x) \leq g(x)$ para todo $x \in D$, entonces*
$$\int_D f(x)dx \leq \int_D g(x)dx.$$
- *La función $|f|$ también es integrable sobre D y*
$$\left| \int_D f(x)dx \right| \leq \int_D |f(x)|dx.$$
- *Si $D = D_1 \cup D_2$, con $D_1 \cap D_2 = \varnothing$ y f es integrable en D_1 y en D_2, entonces*
$$\int_D f(x)dx = \int_{D_1} f(x)dx + \int_{D_2} f(x)dx.$$

Demostración. Supondremos en las propiedades (a)–(d) que $D = R$, un rectángulo. Si no, basta aplicar los resultados obtenidos a las funciones f_D, g_D. Sea S un reticulado de R. Veamos la propiedad (a). Si $R \in S$, entonces:

$$\inf_R f + \inf_R g \leq \inf_R (f+g) \leq \sup_R (f+g) \leq \sup_R f + \sup_R g.$$

Por lo tanto:

$$I_S(f) + I_S(g) \leq I_S(f+g) \leq S_S(f+g) \leq S_S(f) + S_S(g). \tag{1.3}$$

Sean S_n^1 y S_n^2 sucesiones de reticulados tales que:

$$S_{S_n^1}(f) - I_{S_n^1}(f) \to 0, \quad S_{S_n^2}(g) - I_{S_n^2}(g) \to 0.$$

Consideremos un reticulado S_n más fino que, simultáneamente, S_n^1 y S_n^2, por ejemplo el obtenido de las intersecciones de los elementos de ambos. Entonces:

$$S_{S_n}(f) - I_{S_n}(f) \to 0, \quad S_{S_n}(g) - I_{S_n}(g) \to 0.$$

De aquí se sigue:

$$\int_R f = \lim_{n\to\infty} S_{S_n}(f) = \lim_{n\to\infty} I_{S_n}(f), \quad \int_R g = \lim_{n\to\infty} S_{S_n}(g) = \lim_{n\to\infty} I_{S_n}(g),$$

y de las desigualdades (**??**):

$$S_{S_n}(f+g) - I_{S_n}(f+g) \to 0,$$

por lo tanto, $f+g$ es integrable sobre R con:

$$\int_R (f+g) = \lim_{n\to\infty} S_{S_n}(f+g) = \lim_{n\to\infty} I_{S_n}(f+g) = \int_R f + \int_R g.$$

Para probar (b), si $\alpha \geq 0$, entonces:

$$I_S(\alpha f) = \alpha I_S(f), \quad S_S(\alpha f) = \alpha S_S(f).$$

Si $\alpha \leq 0$ entonces:

$$I_S(\alpha f) = \alpha S_S(f), \quad S_S(\alpha f) = \alpha I_S(f).$$

En ambos casos se concluye directamente el resultado.
Para (c), definimos $h(x) := g(x) - f(x)$, que es integrable sobre R por (a) y (b). Entonces:

$$\int_R h(x)dx = \int_R g(x)dx - \int_R f(x)dx.$$

Y como $h(x) \geq 0$ para todo $x \in R$:

$$0 \leq I_S(h) \leq \int_R h(x)dx,$$

lo que implica:

$$\int_R g(x)dx - \int_R f(x)dx \geq 0.$$

Para (d), consideremos un reticulado S y $R \in S$.
- Si $f \geq 0$, entonces $M_R(f) - m_R(f) = M_R(|f|) - m_R(|f|)$.
- Si $f \leq 0$, entonces $M_R(|f|) - m_R(|f|) = M_R(-f) - m_R(-f)$.
- Si f cambia de signo:

$$M_R(|f|) - m_R(|f|) \leq (M_R(f) - m_R(f)) + (M_R(-f) - m_R(-f)).$$

Entonces:

$$S_S(|f|) - I_S(|f|) \leq S_S(f) - I_S(f) + S_S(-f) - I_S(-f).$$

Si S_n es una sucesión de reticulados tal que:

$$S_{S_n}(f) - I_{S_n}(f) \to 0, \quad S_{S_n}(-f) - I_{S_n}(-f) \to 0,$$

entonces:

$$S_{S_n}(|f|) - I_{S_n}(|f|) \to 0,$$

y $|f|$ es integrable. Además, como $\pm f \leq |f|$, se tiene:

$$\left|\int_R f\right| = \left|\int_R (\pm f)\right| \leq \int_R |f(x)|dx.$$

Para (e), si $D = D_1 \cup D_2$ con $D_1 \cap D_2 = \emptyset$, entonces:

$$f_D = f_{D_1} + f_{D_2},$$

y el resultado se sigue por linealidad. $\square$

4.2 Conjuntos Medibles

Ya discutidas las definiciones elementales del concepto de integral, una pregunta fundamental para poder seguir estudiando esto es: ¿qué tipo de regiones D son las apropiadas para calcular integrales? De todas maneras, ya podemos descartar que se pueda en cualquier conjunto. Desarrollemos esta idea con un ejemplo.

Ejemplo 4.2.1. Sea

$$D = \{(x, y) \in \mathbb{Q} \times \mathbb{Q} \ : \ 0 \leq x \leq 1, \ 0 \leq y \leq 1\}.$$

La función $f(x, y) = 1_D(x, y)$ no es integrable, pues sin importar el reticulado S_N que tomemos, sabemos que en cada $R_i^N \in S_N$, por la densidad de los racionales en los reales, siempre obtendremos que

$$M_{R_i^N}(f) = 1, \quad m_{R_i^N}(f) = 0,$$

de manera tal que

$$\mathcal{S}_{S_N}(f) = \sum_{i \in I} M_{R_i^N}(f) \cdot A(R_i^N) = \sum_{i \in I} 1 \cdot A(R_i^N) = \sum_{i \in I} A(R_i^N) = 1,$$

y por otra parte

$$\mathcal{I}_{S_N}(f) = \sum_{i \in I} m_{R_i^N}(f) \cdot A(R_i^N) = \sum_{i \in I} 0 \cdot A(R_i^N) = 0.$$

En consecuencia, dado que

$$\mathcal{S}_{S_N} - \mathcal{I}_{S_N} = 1 \neq 0,$$

concluimos que f no es integrable en D. $\Diamond$

Dado lo recién ilustrado, necesitamos un concepto preliminar para encontrar una clase suficientemente grande de conjuntos donde podamos calcular integrales. Para ello, definamos los siguientes conceptos.

Definición 4.2.2 (Medida Nula)**.** *Decimos que un conjunto $A \subseteq \mathbb{R}^n$ acotado, tiene* medida nula *si para todo $\varepsilon > 0$, existe una colección finita de rectángulos tales que*

$$A \subseteq \bigcup_{i=1}^{m} R_i, \quad \sum_{i=1}^{m} A(R_i) < \varepsilon \ .$$

Definición 4.2.3 (Conjunto Jordan-Medible)**.** *Diremos que un conjunto $D \subseteq \mathbb{R}^n$ es* medible en el sentido de Jordan*, o* Jordan medible*, si $\partial C =$Fr(C) es un conjunto de medida nula.*

A partir de lo anterior, tenemos los siguientes resultados fundamentales.

Proposición 4.2.4. *Consideremos dos funciones continuas $g, h : [a, b] \to \mathbb{R}$ tales que $g(x) \le h(x)$ si $x \in [a, b]$. Entonces, la región del plano*

$$D = \{(x, y) \in \mathbb{R}^2 \ : \ a \le x \le b, \ g(x) \le y \le h(x)\}$$

es Jordan-medible. Se tiene el resultado análogo para funciones $g, h : R \subseteq \mathbb{R}^n \to \mathbb{R}$ y conjuntos $D \subseteq \mathbb{R}^{n+1}$ definidos como

$$D = \{(x, y) \in \mathbb{R}^{n+1} \ : \ x \in R, \ g(x) \le y \le h(x)\}.$$

Juntando todo lo conversado hasta aquí, nuestro resultado final sobre integrabilidad corresponde al siguiente.

Teorema 4.2.5. *Si $D \subseteq \mathbb{R}^n$ es un conjunto Jordan-medible, cerrado y acotado, y además $f : D \to \mathbb{R}$ una función continua. Entonces, f es integrable sobre D. En particular, el volumen de D,*

$$\mathrm{vol}(D) = \int_{\mathbb{R}^n} 1_D \ dx_1 dx_2 \ldots dx_n,$$

está bien definido.

4.3 Cálculo de Integrales

Ya sabemos cómo definir bien la integral, pensando tanto en las funciones que podemos integrar, como en los conjuntos sobre los cuales puedo definir bien esta noción, pero aún no sabemos calcular ninguna.
Esta será nuestra temática en la siguiente sección, donde enunciaremos dos teoremas fundamentales para poder llevar a cabo este tipo de cálculos.

4.3.1 Teorema de Fubini

En las primeras páginas de este capítulo, respondimos la pregunta sobre qué es una integral (definición) y dónde podemos calcularlas (conjuntos medibles).
Ahora responderemos una pregunta igual o más importante, ¿cómo calculamos una integral múltiple? Por ejemplo, si tomamos $R = [0, 3] \times [1, 5]$ y la función $f(x, y) = x^2 y$, busquemos cómo obtener el valor de

$$\iint\limits_R f(x, y) \ dxdy.$$

Idealmente, quisiéramos una herramienta que nos permita hacer esto sin pasar por el cálculo de las sumas superior e inferior para lograrlo. Por suerte la tendremos, pues en la mayoría de los casos podremos usar una técnica iterativa para calcularlas.

Teorema 4.3.1 (Teorema de Fubini)**.** *Dados dos rectángulos $R_1 \in \mathbb{R}^n$ y $R_2 \in \mathbb{R}^m$, definamos el rectángulo producto*

$$R = R_1 \times R_2 = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : x \in R_1, y \in R_2\}.$$

Suponga que $f : R \subset \mathbb{R}^{n+m} \to \mathbb{R}$ es integrable en R, y en donde

$$x \in R_1 \mapsto F_1(x) = \int_{R_2} f(x, y) dy \ , \quad y \in R_2 \mapsto F_2(y) = \int_{R_1} f(x, y) dx \ ,$$

están ambas bien definidas y son integrables donde corresponda. Entonces, si denotamos al vector $z = (x, y) \in \mathbb{R}^{n+m}$, tenemos que

$$\int_R f(z) dz = \int_{R_1} \left(\int_{R_2} f(x, y) dy \right) dx = \int_{R_2} \left(\int_{R_1} f(x, y) dx \right) dy.$$

Nota: E

n términos de notación, si $x \in \mathbb{R}^n$, escribimos

$$\int_R f(x)dx = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_n}^{b_n} f(x_1, x_2, \ldots, x_n)\ dx_1 dx_2 \ldots dx_n.$$

Ejemplo 4.3.2. Retomemos el ejemplo escrito al comienzo, y usemos el Teorema de Fubini para obtener el valor de

$$\iint\limits_{[0,3]\times[1,5]} x^2y\ dxdy = \int\limits_0^3\int\limits_1^5 x^2y\ dxdy.$$

Aplicamos el Teorema de Fubini, integrando primero la variable y y luego la x:

$$\int_0^3\left(\int_1^5 x^2y\ dy\right)dx$$

Calculamos primero la integral interna:

$$\int_1^5 x^2y\ dy = x^2\int_1^5 y\ dy = x^2\left[\frac{y^2}{2}\right]_1^5 = x^2\left(\frac{25}{2}-\frac{1}{2}\right) = x^2\cdot\frac{24}{2} = x^2\cdot 12$$

Ahora resolvemos la integral externa:

$$\int_0^3 12x^2\ dx = 12\int_0^3 x^2\ dx = 12\left[\frac{x^3}{3}\right]_0^3 = 12\cdot\frac{27}{3} = 12\cdot 9 = 108.$$

Si por otro lado, deseábamos integrar primero en la variable x y luego en y, obteníamos:

$$\iint\limits_{[0,3]\times[1,5]} x^2y\ dxdy = \int_1^5\left(\int_0^3 x^2y\ dx\right)dy$$

Calculamos primero la integral interna:

$$\int_0^3 x^2y\ dx = y\int_0^3 x^2\ dx = y\left[\frac{x^3}{3}\right]_0^3 = y\cdot\frac{27}{3} = y\cdot 9$$

Ahora resolvemos la integral externa:

$$\int_1^5 9y\ dy = 9\int_1^5 y\ dy = 9\left[\frac{y^2}{2}\right]_1^5 = 9\left(\frac{25}{2}-\frac{1}{2}\right) = 9\cdot\frac{24}{2} = 9\cdot 12 = 108.$$

$\lozenge$

El Teorema de Fubini también nos permitirá calcular integrales en conjuntos que no son necesariamente rectangulares. Veamos el desarrollo del siguiente ejemplo.

Ejemplo 4.3.3. Calculemos la integral iterada

$$\int_0^8\left(\int_{\sqrt[3]{y}}^2 \sqrt{x^4+1}\ dx\right)dy.$$

En primer lugar, notaoms que la región de integración está definida por:

- $y \in [0,8]$

- Para cada y, $x \in [\sqrt[3]{y}, 2]$

Podemos invertir el orden de integración, notando que

$$x \in [0, 2], \quad \text{y para cada } x, \quad y \in [0, x^3]$$

Una gráfica de la región de integración corresponde a la siguiente

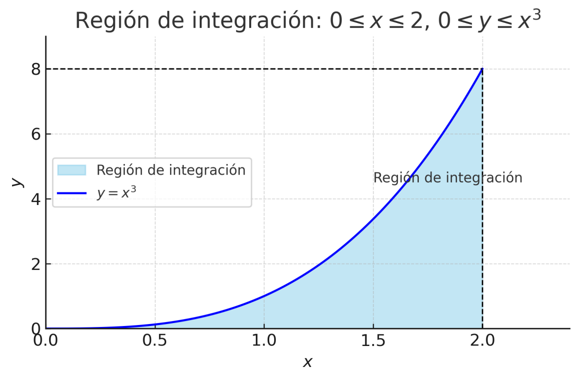


Figura 4.5: Región de Integración

Usamos el Teorema de Fubini, reescribimos la integral como

$$\int_0^8 \left(\int_{\sqrt[3]{y}}^2 \sqrt{x^4+1}\, dx \right) dy = \int_0^2 \left(\int_0^{x^3} \sqrt{x^4+1}\, dy \right) dx.$$

La función $\sqrt{x^4+1}$ no depende de y, así que podemos sacarla de la integral interna:

$$= \int_0^2 \sqrt{x^4+1} \cdot x^3\, dx.$$

Sea $u = x^4 + 1 \Rightarrow du = 4x^3 dx$, entonces:

$$\int_0^2 x^3 \sqrt{x^4+1}\, dx = \frac{1}{4} \int_{u=1}^{17} \sqrt{u}\, du = \frac{1}{4} \cdot \frac{2}{3} \left(17^{3/2} - 1 \right) = \frac{1}{6} \left(17^{3/2} - 1 \right).$$

◊

4.3.2 Teorema del Cambio de Variables

El Teorema del Cambio de Variables nos entrega una herramienta útil de cálculo en caso que la región de integración y/o la función involucrada se puedan expresar más simples al utilizar coordenadas alternativas.

Teorema 4.3.4 (Teorema del Cambio de Variables)**.** *Sea $\Omega \subseteq \mathbb{R}^n$ un conjunto abierto y $T : \Omega \to \mathbb{R}^n$ una función de clase $\mathcal{C}^1$.*
Sea D conjunto abierto y acotado, con $\overline{D} \subseteq \Omega$, y supongamos además que:

- *T es inyectiva en D.*
- *El Jacobiano $DT(u)$ es invertible, para todo $u \in D$.*
- *$T(D)$ es también un conjunto abierto.*

Entonces, para toda función continua $f : \overline{T(D)} \to \mathbb{R}$, se cumple

$$\int\limits_{T(D)} f(x)\ dx = \int\limits_{D} f(T(u))\ |\det(DT(u))|\ du.$$

Existen algunos cambios de variables que son utilizados con gran frecuencia, y por ende vale la pena estudiarlos en sí mismos.

- **Coordenadas polares:** (ρ, θ)
  Estas jugarán el mismo rol que (x, y) en $\mathbb{R}^2$.
- **Coordenadas cilíndricas:** (ρ, θ, z)
  Estas jugarán el mismo rol que (x, y, z) en $\mathbb{R}^3$.
- **Coordenadas esféricas:** (r, θ, φ)
  Estas jugarán el mismo rol que (x, y, z) en $\mathbb{R}^3$.

1. Coordenadas Polares: Escribimos las coordenadas (x, y) en función de las coord. polares (ρ, θ),

$$\begin{pmatrix} x \\ y \end{pmatrix} = T(\rho, \theta) = \begin{pmatrix} \rho \cos\theta \\ \rho \operatorname{sen}\theta \end{pmatrix} \implies \rho = \sqrt{x^2 + y^2},\ \theta = \arctan\left(\frac{y}{x}\right).$$

En este contexto, $\rho \in [0, +\infty)$ y $\theta \in [0, 2\pi)$. Además,

$$\det(DT(\rho, \theta)) = \begin{vmatrix} \cos\theta & -\rho \operatorname{sen}\theta \\ \operatorname{sen}\theta & \rho\cos\theta \end{vmatrix} = \rho\left(\cos^2\theta + \operatorname{sen}^2\theta\right) = \rho.$$

El "elemento de área polar" corresponde a

$$dx\, dy \ \longrightarrow\ \rho\ d\rho\, d\theta.$$

Ejemplo 4.3.5. Calculemos el área de la región dada por

$$C = \{(x, y) \in \mathbb{R}^2\ :\ x^2 + y^2 \leq 4,\ |x| \leq y\},$$

esto es, la integral

$$A = \iint\limits_C 1 dxdy$$

La región está acotada por:

- El círculo $x^2 + y^2 = 4$, es decir, un disco de radio 2 centrado en el origen.
- La desigualdad $|x| \leq y$, o sea, $x \leq y$ y $-x \leq y$, es decir:

  $$y \geq |x|$$

  Esto define la zona **sobre** las rectas $y = x$ y $y = -x$.

Por lo tanto, la región C es la **porción del círculo** entre las rectas $y = x$ y $y = -x$, es decir, un sector circular desde $\theta = \frac{\pi}{4}$ hasta $\theta = \frac{3\pi}{4}$.
Recordando que:

$$x = r\cos\theta, \quad y = r\operatorname{sen}\theta, \quad dxdy = r\, dr\, d\theta$$

La región C en coordenadas polares es:

$$0 \leq r \leq 2, \quad \frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$$

Entonces el área se calcula como:

$$A = \int_{\theta=\frac{\pi}{4}}^{\frac{3\pi}{4}} \int_{r=0}^{2} r\, dr\, d\theta$$

Primero la integral en r:

$$\int_0^2 r\, dr = \left[\frac{r^2}{2}\right]_0^2 = \frac{4}{2} = 2$$

Luego la integral en θ:

$$\int_{\pi/4}^{3\pi/4} 2\, d\theta = 2\left(\frac{3\pi}{4} - \frac{\pi}{4}\right) = 2 \cdot \frac{\pi}{2} = \pi.$$

$\diamondsuit$

2. Coordenadas Cilíndricas: Las coordenadas cartesianas (x, y, z) en función de las coordenadas cilíndricas (ρ, θ, z) quedan como

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = T(\rho, \theta, z) = \begin{pmatrix} \rho\cos\theta \\ \rho\operatorname{sen}\theta \\ z \end{pmatrix} \implies \rho = \sqrt{x^2 + y^2},\ \theta = \arctan\left(\frac{y}{x}\right),\ z = z,$$

En este contexto, $\rho \in [0, +\infty)$, $\theta \in [0, 2\pi)$ y $z \in \mathbb{R}$. Además,

$$\det(DT(\rho, \theta, z)) = \begin{vmatrix} \cos\theta & -\rho\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \rho\left(\cos^2\theta + \operatorname{sen}^2\theta\right) = \rho.$$

El “elemento de volumen” cilíndrico corresponde a

$$dx\, dy\, dz \longrightarrow \rho\, d\rho\, d\theta\, dz.$$

Ejemplo 4.3.6. Usando coordenadas cilíndricas, calculemos el volumen de la región sólida definida por

$$C = \{(x, y, z) \in \mathbb{R}^3\ :\ x^2 + y^2 \leq 9,\ z \geq 1,\ y + z \leq 5\}$$

Recordemos que

$$x = \rho\cos\theta, \quad y = \rho\operatorname{sen}\theta, \quad z = z.$$

Las condiciones que definen al conjunto C, se traducen a:

$$\rho \leq 3,$$

$$z \geq 1,$$

$$z \leq 5 - y = 5 - \rho\operatorname{sen}\theta.$$

De esta manera, obtenemos que C queda dado por los nuevos límites de integración:

$$\rho \in [0, 3], \quad \theta \in [0, 2\pi], \quad z \in [1, 5 - \rho\operatorname{sen}\theta].$$

Dicho esto, el volumen buscado corresponde a

$$V = \iiint_C dV = \int_0^{2\pi} \int_0^3 \int_1^{5-\rho\operatorname{sen}\theta} \rho\, dz\, d\rho\, d\theta,$$

donde el factor ρ corresponde al Jacobiano del cambio a coordenadas cilíndricas.

Si integramos primero respecto a z,

$$\int_1^{5-\rho \operatorname{sen}\theta} \rho\, dz = \rho\,[(5-\rho \operatorname{sen}\theta)-1] = \rho(4-\rho \operatorname{sen}\theta).$$

Entonces,

$$V = \int_0^{2\pi}\int_0^3 \big(4\rho - \rho^2 \operatorname{sen}\theta\big) d\rho\, d\theta.$$

Integrando respecto a ρ,

$$\int_0^3 (4\rho - \rho^2 \operatorname{sen}\theta) d\rho = \int_0^3 4\rho\, d\rho - \operatorname{sen}\theta \int_0^3 \rho^2\, d\rho = 18 - 9\operatorname{sen}\theta.$$

Finalmene, integrando respecto a θ,

$$V = \int_0^{2\pi} \big(18 - 9\operatorname{sen}\theta\big) d\theta = \int_0^{2\pi} 18\, d\theta - 9\int_0^{2\pi} \operatorname{sen}\theta\, d\theta = 36\pi - 0 = 36\pi.$$

$\diamondsuit$

3. Coordenadas Esféricas: Las coordenadas cartesianas, en función de las coord. esféricas (r, θ, φ), son:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = T(r,\theta,\varphi) = \begin{pmatrix} r\cos\theta \operatorname{sen}\varphi \\ r \operatorname{sen}\theta \operatorname{sen}\varphi \\ r\cos\varphi \end{pmatrix}$$

donde $r \in [0, +\infty)$, $\theta \in [0, 2\pi)$ y $\varphi \in [0, \pi]$. Además,

$$\Longrightarrow\; r = \sqrt{x^2+y^2+z^2},\;\; \theta = \arctan\left(\frac{y}{x}\right),\;\; \varphi = \arctan\left(\frac{\sqrt{x^2+y^2}}{z}\right).$$

Entonces, el Jacobiano del cambio de coordenadas es

$$DT(r,\theta,\varphi) = \begin{bmatrix} \cos\theta \operatorname{sen}\varphi & -r\operatorname{sen}\theta \operatorname{sen}\varphi & r\cos\theta\cos\varphi \\ \operatorname{sen}\theta \operatorname{sen}\varphi & r\cos\theta \operatorname{sen}\varphi & r\operatorname{sen}\theta\cos\varphi \\ \cos\varphi & 0 & -r\operatorname{sen}\varphi \end{bmatrix}.$$

Se sigue que, el determinante del Jacobiano es

$$\begin{aligned}
\det(DT(r,\theta,\varphi)) &= \cos\varphi \cdot \det \begin{bmatrix} -r\operatorname{sen}\theta \operatorname{sen}\varphi & r\cos\theta\cos\varphi \\ r\cos\theta \operatorname{sen}\varphi & r\operatorname{sen}\theta\cos\varphi \end{bmatrix} \\
&\quad - 0 + (-r\operatorname{sen}\varphi)\cdot \det \begin{bmatrix} \cos\theta \operatorname{sen}\varphi & -r\operatorname{sen}\theta \operatorname{sen}\varphi \\ \operatorname{sen}\theta \operatorname{sen}\varphi & r\cos\theta \operatorname{sen}\varphi \end{bmatrix} \\
&= \cos\varphi(-r^2\operatorname{sen}^2\theta \operatorname{sen}\varphi\cos\varphi - r^2\cos^2\theta \operatorname{sen}\varphi\cos\varphi) \\
&\quad - (r\operatorname{sen}\varphi)(r\cos^2\theta \operatorname{sen}^2\varphi + r\operatorname{sen}^2\theta \operatorname{sen}^2\varphi) \\
&= -r^2\operatorname{sen}\varphi\cos^2\varphi - r^2\operatorname{sen}\varphi \operatorname{sen}^2\varphi \\
&= -r^2\operatorname{sen}\varphi.
\end{aligned}$$

Luego, al hacer cambio de cariables de cartesianas a esféricas, los "elementos de volumen" diferenciales se reemplazan de la forma

$$dx\,dy\,dz \;\longrightarrow\; r^2\operatorname{sen}\varphi\; dr\,d\theta\,d\varphi.$$

Ejemplo 4.3.7. Usando coordenadas esféricas, calculemos el volumen de la región definida por

$$E = \{(x, y, z) \in \mathbb{R}^3 \ : \ x^2 + y^2 + z^2 \leq 9, \ x \leq y \leq x\sqrt{3}, \ z \geq 0\}$$

Notemos que al reemplazar las coordenadas, obtenemos:

- Por la condición $x^2 + y^2 + z^2 \leq 9$, se tiene $r \leq 3$.
- Por $z \geq 0$, se tiene $r \cos\varphi \geq 0$, es decir, $\varphi \in [0, \pi/2]$.
- La condición $x \leq y \leq x\sqrt{3}$ en el plano xy implica

  $$1 \leq \frac{y}{x} = \tan\theta \leq \sqrt{3},$$

  por lo que

  $$\theta \in \left[\frac{\pi}{4}, \frac{\pi}{3}\right].$$

En consecuencia, el volumen pedido corresponde a:

$$V = \int_{\theta=\pi/4}^{\pi/3} \int_{\varphi=0}^{\pi/2} \int_{r=0}^{3} r^2 \operatorname{sen}\varphi \, dr \, d\varphi \, d\theta.$$

Integrando respecto a r,

$$\int_0^3 r^2 \, dr = \left[\frac{r^3}{3}\right]_0^3 = \frac{27}{3} = 9.$$

Luego, con respecto a φ:

$$\int_0^{\pi/2} \operatorname{sen}\varphi \, d\varphi = [-\cos\varphi]_0^{\pi/2} = 1.$$

Finalmente, integrando respecto a θ:

$$\int_{\pi/4}^{\pi/3} d\theta = \frac{\pi}{3} - \frac{\pi}{4} = \frac{\pi}{12}.$$

En resumidas cuentas,

$$V = 9 \times 1 \times \frac{\pi}{12} = \frac{3\pi}{4}.$$

$\diamondsuit$

4.4 Selección de Problemas

1. Calcule la siguiente integral
$$\int_0^{\pi^2} \int_{\sqrt{y}}^{\pi} \frac{\operatorname{sen}(x^2)}{x^2}\, y^{3/2}\, dxdy.$$

2. Calcule el valor de la integral
$$\iint_P (x-y)^2\ dxdy,$$
donde P es el paralelógramo de vértices $(0,0)$, $(1,1)$, $(2,0)$ y $(1,-1)$.
Indicación: *Puede serle útil bosquejar P para determinar bien sus límites de integración.*

3. Bosqueje la región de integración, y calcule la integral
$$\int_0^1 \int_y^1 \frac{\operatorname{sen}(x)}{x} dxdy.$$

4. *a*) Usando el Teorema de Fubini, determine el valor de la integral
$$I_1 = \int_0^1 \left(\int_y^1 (y-x)\cos(x^3)\ dx \right) dy.$$
b) Usando un cambio de variables adecuado, determine el valor de la integral
$$I_2 = \iiint_D \sqrt{x^2+y^2+z^2}\ dxdydz,$$
donde $D = \{(x,y,z) \in \mathbb{R}^3\ :\ 4 \le x^2+y^2+z^2 \le 25,\ z \ge \sqrt{x^2+y^2}\}$.

5. Calcule el valor de la integral
$$\iiint\limits_E \frac{z(x^2+y^2)}{x^2+y^2+z^2}\ dxdydz,$$
donde E es la intersección del cono de ecuación $z \ge \sqrt{x^2+y^2}$ con la esfera de ecuación $x^2+y^2+z^2 \le 1$. Para ello, proceda como se detalla a continuación:
   a) Bosqueje la región de integración E.
   b) Haciendo un cambio de variables adecuado, reescriba la región E.
   c) Usando el Teorema del Cambio de Variables, calcule la integral propuesta.

6. Calcule el volumen de la región
$$D = \{(x,y,z) \in \mathbb{R}^3\ :\ z \ge \sqrt{x^2+y^2},\ z \le 6 - x^2 - y^2\}.$$

7. Considere la región en $\mathbb{R}^2$ dada por
$$R = \{(x,y) \in \mathbb{R}^2\ :\ (x-1)^2+y^2 \ge 1,\ (x-2)^2+y^2 \le 4\}$$
y considere la función de densidad de masa
$$\rho(x,y) = \frac{y}{x^2+y^2}.$$

En esta pregunta buscamos calcular la masa del cuerpo delimitado por la región R con densidad $\rho(x, y)$, vale decir, el valor de la integral:

$$M = \iint_R \rho(x, y)\ dxdy.$$

Para ello, proceda como se detalla a continuación:

a) Bosqueje la región de integración R.

b) Usando un cambio de variables pertinente, reescriba la región R en sus nuevas variables.

c) Usando el Teorema del Cambio de Variables, calcule el valor de M.

8. *a*) Calcule el valor de

$$I_1 = \iint_C xy\ dxdy,$$

donde $C = \{(x, y) \in \mathbb{R}^2\ :\ x \geq 0,\ y \geq 0,\ x + y \leq 1\}$. Para ello:

1) Bosqueje la región de integración C.
2) Usando el Teorema de Fubini, calcule el valor de I_1.

b) Nuestro objetivo en esta pregunta es calcular el valor de la integral

$$I_2 = \iint_D \frac{4 - x^2 - y^2}{(x^2 + y^2)^{3/2}}\ dxdy,$$

donde $D = \{(x, y) \in \mathbb{R}^2\ :\ x^2 + y^2 \geq 1,\ (x - 1)^2 + y^2 \leq 1\}$. Para ello:

1) Bosqueje la región de integración D.
2) Haciendo un cambio de variables adecuado, reescriba D.
3) Usando el Teorema de Cambio de Variables, calcule el valor de I_2.