

Universidad Autónoma de Madrid

Matemáticas II

Grado en Química

Curso 2024–2025

Roberto Giménez Conejero

Índice

Espacios vectoriales	2
Transformaciones lineales	8
Espacios euclídeos y ortogonalización	12
Ecuaciones diferenciales matriciales de primer orden	19
Funciones de varias variables reales. Cálculo diferencial	24
Funciones de varias variables reales (versión separada)	47
Cálculo diferencial de varias variables (versión separada)	61
Cálculo integral en dos y tres variables	80
Cálculo vectorial: integrales escalares	83
Cálculo vectorial: campos y teoremas	94
Integrales en curvas y superficies	100
Algunos operadores diferenciales importantes	118
Teoremas fundamentales del cálculo vectorial	121

GUÍA DE ESTUDIO: ESPACIOS VECTORIALES.

Matemáticas II

28 de enero de 2025

1. Sistema de ecuaciones lineales

Una *ecuación lineal en las variables* $x_1, \dots, x_n$ *con coeficientes reales* es una ecuación del tipo

$$a_1x_1 + \dots + a_nx_n = b, \quad (1)$$

donde $a_1, \dots, a_n$, b son números reales. Si $b = 0$, la ecuación anterior se dice *ecuación lineal homogénea*. Un *sistema de ecuaciones lineales* es un conjunto finito de ecuaciones del tipo (1). El sistema se dice *homogéneo* si cada ecuación del sistema es una ecuación homogénea. En general, un sistema de m ecuaciones en las variables $x_1, \dots, x_n$ es un sistema del tipo

$$\begin{array}{ccccccc} a_{11}x_1 & + \dots & + a_{1n}x_n & = & b_1 & & \\ \vdots & \dots & \dots & \dots & \vdots & \vdots & \\ a_{m1}x_1 & + \dots & + a_{mn}x_n & = & b_m & & \end{array} \quad (2)$$

El anterior sistema admite una reescritura matricial de la siguiente forma:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}. \quad (3)$$

También lo escribiremos de forma abreviada $A\vec{x} = \vec{b}$. Llamaremos *matriz del sistema* a la matriz A , y *matriz aumentada del sistema* a la matriz $(A|\vec{b})$.

Una n -upla $(s_1, \dots, s_n)$ de números reales se dice *una solución del sistema* (2) si satisface el sistema, es decir $A\vec{s} = \vec{b}$, donde $\vec{s} := (s_1, \dots, s_n)^t$. Diremos que el sistema (2) es un sistema *compatible determinado* si tiene una única solución. Será *compatible indeterminado* si tiene más de una solución. Caso de no tener solución el sistema se dice *incompatible*.

1.1. Algoritmo de Gauss-Jordan para la resolución de sistemas de ecuaciones lineales

La resolución de sistemas de ecuaciones lineales es una herramienta básica en la resolución de problemas tanto de ciencia básica como de ingeniería. El método aplicado para resolverlos es el dado por el algoritmo de Gauss-Jordan. A continuación expondremos los pasos a seguir, si bien son necesarios algunos conceptos previos

1.1.1. Matrices escalonadas y transformaciones elementales

Fijada una matriz M , diremos que su *fila* i -ésima es *nula* si todas las entradas de esta fila son 0. Una matriz M se dice escalonada si la fila i -ésima tiene más ceros a la izquierda que la fila j -ésima para $j < i$. Por ejemplo, la siguiente matriz está en forma escalonada:

$$M = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 5 & 10 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (4)$$

El sistema lineal en variables x, y, z asociado a esta matriz sería:

$$x + 2z = 0, \quad 5z = 10. \quad (5)$$

El primer elemento no nulo de una fila dada se llama *elemento pivote*. La columna en la que hay un elemento pivote se llama *columna pivote*. Una matriz está en forma *escalonada reducida* es una matriz escalonada, los elementos pivote son todos unos, a los que llamaremos *unos pivote del sistema*, y los únicos elementos no nulos en las columnas pivote son los unos pivote.

Todas las matrices, cuadradas o no, se pueden escalar usando las siguientes transformaciones, llamadas *transformaciones elementales*:

1. Intercambiar dos filas.
2. Multiplicar una fila por una constante c no nula.
3. Multiplicar una fila por una constante $c \neq 0$ y sumarsela a otra.

Ilustraremos estas transformaciones sobre la matriz M dada en (4) cuyas filas serán denotadas en números romanos por I, II, III .

$$M \xrightarrow{(1/5)II} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{(-2)II+I} E = \begin{pmatrix} 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (6)$$

Es de señalar que E es la matriz escalonada reducida asociada a M . El sistema lineal en variables x, y, z asociado a esta matriz E sería:

$$x = -4, \quad z = 2. \quad (7)$$

Observa que tanto (5) como (7) tiene por espacio de soluciones:

$$L := \{(-4, t, 2) \mid t \in \mathbb{R}\} = (-4, 0, 2) + (0, 1, 0)\mathbb{R}. \quad (8)$$

Por tanto, L es la recta de $\mathbb{R}^3$ que pasa por el punto $(-4, 0, 2)$ y tiene vector director $\vec{u} = (0, 1, 0)$.

El procedimiento que permite obtener una matriz escalonada reducida partiendo una matriz ampliada de un sistema lineal mediante transformaciones elementales, se llama *Algoritmo de Gauss-Jordan*. Lo describimos a continuación.

1.1.2. Algoritmo de Gauss-Jordan

El algoritmo de Gauss consta de los siguientes pasos:

- **Input:** Matriz ampliada del sistema, $(A|\vec{b})$ de tamaño $m \times (n + 1)$.
 - **Output:** Matriz escalonada $(C|\vec{d})$
1. Localiza la primera columna no nula de la matriz $(A|\vec{b})$. Supongamos que es la columna j .
 2. Si a_{ij} es el primer elemento no nulo de la columna j , entonces
 - Intercambia fila 1 por fila j y divídela por a_{ij} . Has obtenido así el primer 1 pivote del sistema.
 - Mediante las transformaciones elementales 1, 2, 3, procede a eliminar los elementos no nulos de la columna j . Así obtendrás una columna j en el que el primer elemento es un 1 y todos los demás son 0.
 3. Considera la matriz obtenida al eliminar la primera fila de la matriz anterior. Denotemos esta matriz por $(A_1|\vec{b}_1)$.
 4. Repite el proceso anterior para la matriz $(A_1|\vec{b}_1)$, es decir ve al Paso 1, hasta que $A_1 = \mathbf{0}$ o hayas realizado m pasos del algoritmo.

Llamaremos *Algoritmo de Gauss-Jordan* para la matriz $(A|\vec{b})$ al algoritmo de Gauss aplicado a la matriz escalonada $(C|\vec{d})$ cuando en esta matriz se consideran las filas numeradas en orden inverso. Su aplicación a la matriz $(A|\vec{b})$ permite obtener la forma escalonada reducida $(E|\vec{f})$ de la matriz ampliada $(A|\vec{b})$. Un ejemplo numérico puede verse en (6).

2. Espacios vectoriales

2.1. Espacio de soluciones de un sistema lineal

Consideraremos en este apartado un sistema lineal de la forma $A\vec{x} = \vec{b}$ como en (3), es decir con m ecuaciones y n incógnitas. Su espacio de soluciones S es un subconjunto de $\mathbb{R}^n$ cuya estructura algebraica estudiaremos en esta sección. Distinguiremos dos comportamientos distintos de este conjunto según el sistema lineal sea homogéneo o no.

Caso homogéneo. El espacio de soluciones S de un sistema homogéneo tiene las siguientes propiedades

1. El vector $\vec{0}$ siempre está en S
2. Si $\vec{u}, \vec{v}$ están en S , entonces su suma $\vec{u} + \vec{v}$ está en S .
3. Si c es una constante real y $\vec{u}$ está en S , entonces $c\vec{u}$ está en S .

Si un subconjunto de $\mathbb{R}^n$ satisface las propiedades anteriores diremos que es un *espacio vectorial en $\mathbb{R}^n$* . En particular, el espacio de soluciones de un sistema lineal homogéneo es un subespacio vectorial de $\mathbb{R}^n$.

Caso no homogéneo. En primer lugar, para describir el espacio de soluciones de un sistema no homogéneo $A\vec{x} = \vec{b}$, tenemos que considerar el espacio S_h de soluciones del sistema homogéneo asociado:

$$A\vec{x} = \vec{0}.$$

Entonces, el espacio de soluciones S del sistema no homogéneo tiene las siguientes propiedades

1. El vector $\vec{0}$ nunca está en S
2. Si p_0 es una solución del sistema no homogéneo, y $\vec{u}$ está en S_h , entonces $p_0 + \vec{u}$ está en S . Es decir, $S = p_0 + S_h$.
3. Si p, q están en S , entonces el vector $\vec{pq}$ está en S_h .

Esta estructura algebraica que posee S la llamaremos *estructura afín*, y a S *espacio afín asociado al sistema $A\vec{x} = \vec{b}$* . El espacio de soluciones del sistema homogéneo asociado recibe el nombre de *espacio de direcciones de S* .

2.2. Espacios vectoriales generales

Un *espacio vectorial real* es un conjunto V no vacío de objetos llamados *vectores* en el que están definidas dos operaciones, llamadas *suma*, denotada por $+$ y *producto por un escalar*, denotado por $\cdot$, que satisfacen las siguientes propiedades:

- E. 1. Si $\vec{u}, \vec{v}$ están en V , entonces $\vec{u} + \vec{v}$ está en V .
- E. 2. Si $\vec{u}, \vec{v}$ están en V , entonces $\vec{u} + \vec{v} = \vec{v} + \vec{u}$.
- E. 3. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.
- E. 4. Existe un elemento $\vec{0}$, llamado *vector cero*, tal que $\vec{u} + \vec{0} = \vec{u}$.
- E. 5. Para cada vector $\vec{u}$ existe un vector, llamado *su vector opuesto* y denotado por $-\vec{u}$ tal que $\vec{u} + (-\vec{u}) = \vec{0}$.
- E. 6. Si c es un escalar real y $\vec{u}$ un vector, entonces $c \cdot \vec{u}$ es un vector de V .
- E. 7. Si c es un escalar real y $\vec{u}, \vec{v}$ están en V , entonces $c \cdot (\vec{u} + \vec{v}) = c \cdot \vec{u} + c \cdot \vec{v}$.
- E. 8. Si c, d son escalares reales y $\vec{u}$ un vector de V , entonces $(c + d) \cdot \vec{u} = c \cdot \vec{u} + d \cdot \vec{u}$.
- E. 9. Si c, d son escalares reales y $\vec{u}$ un vector de V , entonces $(cd) \cdot \vec{u} = c \cdot (d \cdot \vec{u})$.
- E. 10. Para cualquier vector $\vec{u}$ de V , se tiene que $1 \cdot \vec{u} = \vec{u}$.

Como consecuencia de las propiedades anteriores se da también para vector $\vec{u}$ en V :

1. $0 \cdot \vec{u} = \vec{0}$.
2. $c \cdot \vec{0} = \vec{0}$.
3. $-\vec{u} = (-1) \cdot \vec{u}$.

Ejemplo 1. Listaremos a continuación varios ejemplos de espacios vectoriales.

a) El espacio de soluciones de un sistema de ecuaciones lineales homogéneo.

b) $V = \mathbb{R}^n$ con las operaciones:

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n), \quad (9)$$

$$c \cdot (a_1, \dots, a_n) = (ca_1, \dots, ca_n). \quad (10)$$

c) $V = \{a_0 + a_1x + a_2x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\}$, el espacio de polinomios con coeficientes reales de grado menor o igual que 2.

d) $V = \mathbf{M}_{m,n}(\mathbb{R})$, el espacio de matrices $m \times n$ con entradas reales.

e) $V = \mathcal{C}(D, \mathbb{R})$, el espacio de funciones continuas $f : D \rightarrow \mathbb{R}$ definida en el dominio D .

f) El espacio de sucesiones números reales es un espacio vectorial.

2.2.1. Subespacios vectoriales

Consideremos en este apartado un espacio vectorial real V y W un subconjunto no vacío de V . Diremos que W es un *subespacio vectorial* de V , si en W se satisfacen las propiedades:

1. El vector $\vec{0}$ siempre está en W
2. Si $\vec{u}, \vec{v}$ están en W , entonces su suma $\vec{u} + \vec{v}$ está en W .
3. Si c es una constante real y $\vec{u}$ está en W , entonces $c\vec{u}$ está en W .

Ejemplo 2. En $\mathbb{R}^3$ los siguientes subconjuntos son subespacios vectoriales:

- $W_1 = \{(x, y, z) \mid x + y + z = 0\}$.
- $W_2 = \{(-2t, t, t) \mid t \in \mathbb{R}\}$.

En algunas ocasiones tenemos que considerar espacios vectoriales generados por una familia de vectores. Estos espacios admiten una descripción matemática precisa. La detallamos a continuación.

Definición 1. Dados unos vectores $\vec{v}_1, \dots, \vec{v}_\ell$ de un espacio vectorial V , denotaremos por $\langle \vec{v}_1, \dots, \vec{v}_\ell \rangle$ al conjunto (se muestran ahora varias notaciones de lo mismo)

$$\langle \vec{v}_1, \dots, \vec{v}_\ell \rangle = \langle \{\vec{v}_1, \dots, \vec{v}_\ell\} \rangle = \text{Span}_{\mathbb{R}}\{\vec{v}_1, \dots, \vec{v}_\ell\} := \{c_1\vec{v}_1 + \dots + c_\ell\vec{v}_\ell \mid c_j \in \mathbb{R}\}. \quad (11)$$

Este subespacio vectorial de V se llama el *subespacio vectorial generado* por $\{\vec{v}_1, \dots, \vec{v}_\ell\}$.

Ejemplos de estos subespacios son las rectas y planos de $\mathbb{R}^3$ que pasan por el origen.

Observación 2. La familia de subespacios vectoriales de un espacio vectorial tiene algunas propiedades algebraicas importantes. Por ejemplo

2.3. Dependencia lineal y bases. Dimensión

En un espacio vectorial V pueden existir relaciones entre sus elementos de varios tipos. En esta sección estudiaremos las relaciones (lineales) entre los vectores.

Diremos que $\vec{v}_1, \dots, \vec{v}_\ell$ son *linealmente independientes* si cada combinación lineal del tipo $c_1\vec{v}_1 + \dots + c_\ell\vec{v}_\ell = \vec{0}$ solo se da si los escalares son nulos, es decir $c_1 = 0, \dots, c_\ell = 0$. En otro caso diremos que son linealmente dependientes.

Si el espacio vectorial V está generado por una familia finita de vectores $\vec{v}_1, \dots, \vec{v}_d$ que son linealmente independientes, entonces decimos que el conjunto $B = \{\vec{v}_1, \dots, \vec{v}_d\}$ son una *base* de V . El número d se llama la *dimensión* de V , y se escribe $\dim V = d$.

Ejemplo 3. 1. El espacio vectorial $\mathbb{R}^3$ tiene dimensión 3.

2. El espacio de matrices 2×3 tiene dimensión 6.

3. El espacio de los polinomios de grado ≤ 2 tiene dimensión 3.

2.4. Coordenadas y bases

Fijemos un espacio vectorial V de dimensión d , y una base $B = \{\vec{v}_1, \dots, \vec{v}_d\}$. Consideremos un vector arbitrario $\vec{w} \in V$. Entonces este vector se puede expresar como combinación lineal de los vectores de B ya que $V = \langle B \rangle$. En concreto,

$$\vec{w} = \lambda_1\vec{v}_1 + \dots + \lambda_d\vec{v}_d. \quad (12)$$

La d -upla $(\lambda_1, \dots, \lambda_d)$ se llama las *coordenadas de en la base B* .

Es de señalar que tupla depende de la base. Es decir, si cambiamos la base a otra B' , cambian las coordenada que identifican al mismo vector $\vec{w}$. Las fórmulas que relacionan las coordenadas en la base B con las coordenadas en la base B' se llaman *fórmulas de cambio de coordenadas de la base B a la base B'* . Ilustraremos este hecho con un ejemplo.

Ejemplo 4. Consideremos la base canónica de $\mathbb{R}^2$, $B = \{\vec{e}_1 = (1, 0), \vec{e}_2 = (0, 1)\}$, y la base $B' = \{\vec{u}_1 = \vec{e}_1 + \vec{e}_2, \vec{u}_2 = \vec{e}_1 - \vec{e}_2\}$.

Entonces, fijado un vector $\vec{w}$ en $\mathbb{R}^2$, se tiene:

$$\vec{w} = \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 = \lambda'_1\vec{u}_1 + \lambda'_2\vec{u}_2.$$

Entonces,

$$\vec{w} = \lambda'_1(\vec{e}_1 + \vec{e}_2) + \lambda'_2(\vec{e}_1 - \vec{e}_2) = (\lambda'_1 + \lambda'_2)\vec{e}_1 + (\lambda'_1 - \lambda'_2)\vec{e}_2.$$

En consecuencia, obtenemos la siguiente fórmula de cambio de base de B a B' :

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda'_1 \\ \lambda'_2 \end{pmatrix}.$$

GUÍA DE ESTUDIO:

TEMA2: TRANSFORMACIONES LINEALES

Prof. Roberto Giménez Conejero

3 de febrero de 2025

1. Transformaciones lineales. Ecuación matricial

En esta sección consideraremos dos espacio vectoriales reales V_1 y V_2 y una aplicación entre ellos, $f : V_1 \rightarrow V_2$. Diremos que f es una *aplicación lineal* si verifica las siguientes propiedades:

$$1. f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}).$$

$$2. f(\lambda \vec{u}) = \lambda f(\vec{u}).$$

para $\vec{u}, \vec{v}$ en V_1 y toda constante real λ .

Fijado un espacio vectorial V de dimensión d , podemos considerar sólomente las aplicaciones lineales de V en V . Fijaremos para estas aplicaciones lineales $f : V \rightarrow V$ una base $B = \{\vec{v}_1, \dots, \vec{v}_d\}$ en V para estudiarlas. Entonces, si conocemos las imágenes de los vectores de la base, es decir

$$f(\vec{v}_i) = a_{i1}\vec{v}_1 + \dots + a_{id}\vec{v}_d, \quad i = 1, \dots, d, \quad (1)$$

La *ecuación matricial* de f está dada por

$$\begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1d} \\ \vdots & \dots & \vdots \\ a_{d1} & \dots & a_{dd} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \quad \text{o también } \vec{y} = A\vec{x}. \quad (2)$$

El *núcleo* de f , denotado por $Nuc(f)$, es el espacio de soluciones del sistema homogéneo $A\vec{x} = \vec{0}$, la *imagen* de f , denotada por $Im(f)$, es el espacio vectorial generado por las columnas de la matriz A . Ilustraremos este concepto con ejemplos.

Ejemplo 1. 1. La aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$f(x, y) = (2x + y, x - y) = (x', y'),$$

es una aplicación lineal en $\mathbb{R}^2$. Tiene por ecuación matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Esta transformación tiene núcleo $Nuc(f) = \{\vec{0}\}$ e imagen $Im(f) = \mathbb{R}^2$.

2. El giro de $\pi/4$ en $\mathbb{R}^2$ viene dado por

$$f(x, y) = \left(\frac{1}{\sqrt{2}}x - \frac{1}{\sqrt{2}}y, \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y \right) = (x', y').$$

Tiene por ecuación matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Los ejemplos anteriores sugieren que la ecuación matricial de una transformación $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ podría depender de la base B elegida. Esto es cierto, y la fórmula que relaciona la escritura matricial en la base B con otra en otra base B' es la siguiente:

$$M = P^{-1}MP, \quad (3)$$

donde P es la matriz de cambio de la base B a la base B' . Verificaremos esa fórmula sobre un ejemplo.

Ejemplo 2. Consideremos la base canónica de $\mathbb{R}^2$, $B = \{\vec{e}_1 = (1, 0), \vec{e}_2 = (0, 1)\}$, y la base $B' = \{\vec{u}_1 = \vec{e}_1 + \vec{e}_2, \vec{u}_2 = \vec{e}_1 - \vec{e}_2\}$. Entonces, fijado un vector $\vec{w}$ en $\mathbb{R}^2$, se tiene:

$$\vec{w} = \lambda_1 \vec{e}_1 + \lambda_2 \vec{e}_2 = \lambda'_1 \vec{u}_1 + \lambda'_2 \vec{u}_2.$$

En consecuencia, obtenemos la siguiente fórmula de cambio de base de B a B' :

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda'_1 \\ \lambda'_2 \end{pmatrix}, \quad P := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Luego, la transformación f del ejemplo 1, tiene por ecuación en la base B' :

$$\begin{pmatrix} \tilde{\lambda}'_1 \\ \tilde{\lambda}'_2 \end{pmatrix} = P^{-1} \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} P \begin{pmatrix} \lambda'_1 \\ \lambda'_2 \end{pmatrix} = \begin{pmatrix} 3/2 & 1/2 \\ 3/2 & -3/2 \end{pmatrix} \begin{pmatrix} \lambda'_1 \\ \lambda'_2 \end{pmatrix}.$$

En consecuencia en la base B' f tiene la fórmula:

$$f(\lambda'_1, \lambda'_2) = \left(\frac{3}{2}\lambda'_1 + \frac{1}{2}\lambda'_2, \frac{3}{2}\lambda'_1 - \frac{3}{2}\lambda'_2 \right).$$

2. Autovalores y autovectores

Fijemos en este apartado una transformación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y trataremos de estudiar aquellas direcciones para las cuales la transformación se comporta como una dilatación en dicha dirección. Fijemos una base inicial B en la que f tiene por matriz M .

Un vector $\vec{v}$ de $\mathbb{R}^n$ se dice *autovector* o también *vector propio* de f si para un cierto escalar λ se tiene que $f(\vec{v}) = \lambda\vec{v}$. Al correspondiente escalar se le llama *autovalor* o *valor propio* de f .

Para calcular los posibles autovalores y autovectores de f utilizaremos el siguiente procedimiento:

1. Construimos la matriz $M - \lambda I_n$.
2. Calculamos las raíces reales del polinomio $p_f(\lambda) := \det(M - \lambda I_n)$, $\lambda_1, \dots, \lambda_r$, y sus multiplicidades $m_1, \dots, m_r$ (llamadas *multiplicidades algebraicas*).

3. Para cada $i = 1, \dots, r$, se define el subespacio vectorial

$$E(\lambda_i) := \{\vec{v} \mid (M - \lambda_i I_n)\vec{v} = 0\}, \quad (4)$$

llamado *el autoespacio asociado al autovalor λ_i* , o también *el espacio propio asociado al autovalor λ_i* .

4. Para cada $i = 1, \dots, r$, se calcula una base B_i de $E(\lambda_i)$, y se define $d_i := \dim E(\lambda_i)$ (llamadas *multiplicidades geométricas*). Todos los elementos de $\cup_i E(\lambda_i)$ son autovectores de f .

El polinomio $p_f(\lambda)$ se llama *polinomio característico de f* . Cada λ_i encontrado mediante el proceso anterior es un autovalor de f .

Aplicando el proceso a los ejemplos anteriores, se observa que en el ejemplo 1.1 los autovalores son

$$\lambda_1 = \frac{1}{2} + \frac{1}{2} \sqrt{13}, \quad \lambda_2 = \frac{1}{2} - \frac{1}{2} \sqrt{13},$$

cuyos correspondientes autoespacios son las rectas

$$E(\lambda_1) = \langle (-2, 3 - \sqrt{13}) \rangle, \quad E(\lambda_2) = \langle (-2, 3 + \sqrt{13}) \rangle.$$

Por otro lado, el ejemplo 1.2 no tiene autovalores, y por tanto tampoco autoespacios asociados.

3. Diagonalización de matrices

Una transformación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y diremos que es *diagonalizable* si existe una base B en la la transformación se comporta como una dilatación en cada dirección correspondiente a cada vector de esta base. Es decir los enunciados siguientes son equivalentes:

1. f es diagonalizable
2. $\mathbb{R}^n$ tiene una base formada por autovectores de f .

El siguiente resultado lo usaremos para garantizar la posibilidad de diagonalizar f . Con las notaciones de la sección 2, enunciamos el criterio.

Teorema 1. *Fijada la transformación $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, si aplicado el procedimiento cálculo de autovalores y autovectores de la sección 2, se tiene que $m_i = d_i$, para $i = 1, \dots, r$, entonces f es diagonalizable. Además la matriz de f en la base $B := \cup_i B_i$ es una matriz diagonal.*

Observación 2. *Observa que, en particular, si f tiene n autovalores distintos, entonces f es diagonalizable.*

Es de señalar que el ejemplo 1.1 da una transformación lineal diagonalizable, y que la transformación del ejemplo 1.2 no es diagonalizable.

3.1. Matrices simétricas y su diagonalización

En esta sección comprobaremos sobre ejemplos el siguiente resultado general.

Teorema 3. *Toda matriz cuadrada simétrica es diagonalizable.*

Caso de matrices 2×2 . Consideremos una aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por una matriz simétrica

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Su polinomio característica es

$$p_f(\lambda) = \det(A - \lambda I_2) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{12} & a_{22} - \lambda \end{pmatrix} = \lambda^2 - (a_{11} + a_{22})\lambda + \det(A).$$

Este polinomio de grado 2 en λ siempre tiene sus raíces reales ya que su discriminante es positivo, como muestra el siguiente cálculo:

$$\Delta := (a_{11} + a_{22})^2 - 4 \cdot 1 \det(A) = a_{11}^2 + a_{22}^2 + 2a_{11}a_{22} - 4(a_{11}a_{22} - a_{12}^2) = (a_{11} - a_{22})^2 + 4a_{12}^2.$$

Ilustraremos el teorema 3 para el caso de matrices 3×3 con el siguiente ejemplo.

Ejemplo 3. *La siguiente matriz:*

$$A := \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

tiene por polinomio característico $p_A(\lambda) = \lambda^3 - 6\lambda^2 + 7\lambda - 2$ cuyas raíces son

$$1, \quad \frac{5}{2} - \frac{1}{2}\sqrt{17}, \quad \frac{5}{2} + \frac{1}{2}\sqrt{17}.$$

Luego, por la nota 2, es diagonalizable.

GUÍA DE ESTUDIO: ESPACIOS EUCLÍDEOS Y ORTOGONALIZACIÓN

Prof. Roberto Giménez Conejero

10 de febrero de 2025

1. Espacios euclídeos

Un *espacio euclídeo* es un espacio vectorial real E dotado de un producto escalar, es decir, de una regla que asigne a cada par de vectores $\vec{u}, \vec{v}$ un número real designado por $\langle \vec{u}, \vec{v} \rangle$, de forma que se cumplen las siguientes propiedades:

1. Propiedad simétrica: $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.
2. Propiedad distributiva: $\langle \vec{u} + \vec{w}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{w}, \vec{v} \rangle$.
3. $\langle \lambda \vec{u}, \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle$ para cada $\lambda \in \mathbb{R}$.
4. $\langle \vec{u}, \vec{u} \rangle \geq 0$, y sólo es 0 si $\vec{u} = \vec{0}$.

En inglés se llama *scalar product* o *dot product*, porque a veces también se denota como

$$\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}.$$

Se define el módulo o longitud (o norma) de un vector $\vec{u} \in E$ como

$$\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}. \quad (1)$$

Ejemplo 1.1. 1. Si $E = \mathbb{R}^2$, y $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, entonces la fórmula

$$\langle \vec{u}, \vec{v} \rangle := x_1 x_2 + y_1 y_2 \quad (2)$$

define un producto escalar en E .

2. Si $E = \mathbb{R}^2$, y $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, entonces la fórmula

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (3)$$

define un producto escalar en E . En este caso, el módulo de un vector tiene la fórmula

$$\|\vec{u}\| = \sqrt{x_1^2 + 2x_1 y_1 + 2y_1^2} = \sqrt{(x_1 + y_1)^2 + y_1^2}.$$

3. Si E es el espacio de funciones continuas reales en el intervalo $[a, b]$, denotado por $C([a, b])$, se define para cada par de funciones f, g en E el producto escalar

$$\langle f, g \rangle := \int_a^b f(t)g(t) dt, \quad (4)$$

obteniéndose así un espacio euclídeo de funciones.

En un espacio euclídeo se cumple la siguiente desigualdad.

Proposición 1 (Desigualdad de Schwarz). *En un espacio euclídeo E , se da*

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|,$$

para cualesquiera $\vec{u}, \vec{v}$ en E .

Demostración. Consideremos dos vectores cualesquiera $\vec{u}, \vec{v}$ en E , y λ un escalar real arbitrario. Entonces,

$$\langle \lambda \vec{u} - \vec{v}, \lambda \vec{u} - \vec{v} \rangle \geq 0, \quad (5)$$

por las propiedades del producto escalar. Desarrollando la desigualdad anterior se obtiene

$$\lambda^2 \|\vec{u}\|^2 - 2 \langle \vec{u}, \vec{v} \rangle + \|\vec{v}\|^2 \geq 0.$$

La parte izquierda de esta ecuación es un polinomio cuadrático en λ que no puede tener raíces reales y distintas. En consecuencia su discriminante ha de ser siempre ≤ 0 . Por tanto $\langle \vec{u}, \vec{v} \rangle^2 - \|\vec{u}\|^2 \|\vec{v}\|^2 \leq 0$. Tomando raíces cuadradas obtenemos $|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$, que es la desigualdad pedida. $\square$

Consecuencia de la desigualdad de Cauchy-Schwarz es que en los espacios vectoriales euclídeos se puede establecer un concepto de ángulos entre sus elementos. Por tanto podremos hablar de elementos ortogonales, concepto análogo al de vectores perpendiculares en el espacio euclídeo $\mathbb{R}^2$ con la métrica usual.

Ortogonalidad. Sea E un espacio euclídeo con producto escalar $\langle, \rangle$. Dados dos vectores $\vec{u}, \vec{v}$ en E , diremos que son ortogonales (respecto de $\langle, \rangle$) si $\langle \vec{u}, \vec{v} \rangle = 0$. Además si W es un subespacio de E , se define su complemento ortogonal $W^\perp$ como el conjunto

$$W^\perp = \{ \vec{w} \in E \mid \langle \vec{w}, \vec{u} \rangle = 0 \text{ para todo } \vec{u} \in W \}. \quad (6)$$

En estos espacios existe una noción de *ángulo* entre vectores. En concreto, si $\vec{u}, \vec{v}$ están en E , definimos el *coseno del ángulo* que forman $\vec{u}$ y $\vec{v}$ como

$$\cos \angle(\vec{u}, \vec{v}) := \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \cdot \|\vec{v}\|}. \quad (7)$$

En consecuencia, el *ángulo* que forman $\vec{u}$ y $\vec{v}$ es

$$\angle(\vec{u}, \vec{v}) := \arccos \left(\frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \cdot \|\vec{v}\|} \right) \in [0, \pi]. \quad (8)$$

Ejemplo 1.2. Si $E = \mathbb{R}^2$, y consideramos el producto escalar definido por la fórmula

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (9)$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E , entonces

$$\angle((1, 0), (0, 1)) = \arccos \left(\frac{1}{1 \cdot \sqrt{2}} \right) = \frac{\pi}{4}. \quad (10)$$

En los espacios euclídeos se verifica el Teorema de Pitágoras generalizado.

Teorema 2 (Teorema de Pitágoras generalizado). *Sea E un espacio euclídeo, y $\vec{u}, \vec{v}$ dos vectores ortogonales de E , entonces:*

$$\|\vec{u} + \vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2. \quad (11)$$

Definición 3. *Dos subespacios W_1 y W_2 de un espacio euclídeo E se dicen ortogonales, y se escribe $W_1 \perp W_2$, si todos los vectores de W_1 son ortogonales a todos los vectores de W_2 .*

En general, en un espacio euclídeo todo subespacio W de E tiene asociado otro formado por los vectores que son ortogonales a los de W , se llama *el complemento ortogonal de W* , y se denota por $W^\perp$. Es decir,

$$W^\perp := \{\vec{y} \in E \mid \vec{y} \perp \vec{x}, \text{ para cada vector } \vec{x} \in W\} \quad (12)$$

Además, se dan las siguientes propiedades:

- A. $W_1 \perp W_2$ si y sólo si todos los vectores de una base de W_1 son ortogonales a todos los vectores de una base de W_2 .
- B. Si $W_1 \perp W_2$, entonces $W_1 \cap W_2 = \{\vec{0}\}$.
- C. Se tiene que $E = W + W^\perp$.

Ejemplo 1.3. *El complemento ortogonal del subespacio $W = L(\vec{u}, \vec{v})$, con $\vec{u} = (1, 0, 1)$, $\vec{v} = (2, -1, 1)$ de $\mathbb{R}^3$ cuando se considera el producto escalar habitual es la recta $W^\perp$ de vector director $\vec{r} = (1, 1, -1)$. Se tiene que $\mathbb{R}^3 = W + W^\perp$. Entonces, cada vector de $\mathbb{R}^3$ se pone de forma única como suma de dos vectores, uno que está en W (a veces llamada la componente horizontal del vector) y otro en $W^\perp$ (a veces llamada la componente vertical del vector). Por ejemplo:*

$$(7, 2, 2) = (4, -1, 3) + (3, 3, -1) = 2 \cdot (1, 0, 1) + (2, -1, 1) + 3 \cdot (1, 1, -1).$$

1.1. Espacios afines euclídeos. Distancia

En muchas aplicaciones el espacio vectorial $E = \mathbb{R}^n$ se considera como espacio de puntos $A = \mathbb{R}^n$ y no como espacio de vectores. Esto es debido a que no se establece a priori ninguna referencia cartesiana para calcular. En estas situaciones se dice que consideramos a $\mathbb{R}^n$ como espacio afín A asociado al espacio vectorial euclídeo $E = \mathbb{R}^n$. Por esta situación podemos definir el vector que une dos puntos $P = (p_1, \dots, p_n)$, $Q = (q_1, \dots, q_n)$ como

$$\vec{PQ} := (q_1 - p_1, \dots, q_n - p_n). \quad (13)$$

También podemos definir una distancia entre puntos con la fórmula siguiente:

$$d(P, Q) := \|\vec{PQ}\|. \quad (14)$$

Ilustraremos estos conceptos sobre un ejemplo.

Ejemplo 1.4. *Sea $E = \mathbb{R}^2$, con el producto escalar definido por la fórmula*

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (15)$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E . Los puntos $P = (1, 1)$ y $Q = (1, 0)$ en $A = \mathbb{R}^2$ definen el vector $\vec{PQ} = (0, -1)$ cuyo módulo es $\sqrt{2}$. En consecuencia, para este producto escalar, se tiene que $d(P, Q) = \sqrt{2}$.

2. Ortogonalidad. Proceso de ortonormalización

Fijaremos en esta sección un espacio vectorial euclídeo E de dimensión n con producto escalar $\langle, \rangle$. En E una base se dice *ortogonal* si sus elementos son ortogonales dos a dos; si además, los elementos de una base ortogonal son de longitud 1, la base se dice *ortonormal*. Es decir, una base $B = \{\vec{u}_1, \dots, \vec{u}_n\}$ es ortonormal si se cumple que

$$\langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Si $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ es una base ortonormal de E , y $\vec{u}$ es un vector, entonces se da la siguiente igualdad

$$\vec{u} = \langle \vec{u}, \vec{e}_1 \rangle \vec{e}_1 + \dots + \langle \vec{u}, \vec{e}_n \rangle \vec{e}_n, \quad (16)$$

es decir, las coordenadas $(\lambda_1, \dots, \lambda_n)$ de $\vec{u}$ en B se calculan según la fórmula:

$$\lambda_1 = \langle \vec{u}, \vec{e}_1 \rangle, \dots, \lambda_n = \langle \vec{u}, \vec{e}_n \rangle. \quad (17)$$

Además, en este espacio siempre se puede construir una base ortogonal a partir de una dada; es el método (o algoritmo) de ortogonalización de Gram-Schmidt. Se fundamenta en el siguiente teorema.

Teorema 4. Sean $\vec{u}_1, \dots, \vec{u}_n$ vectores linealmente independientes en un espacio euclídeo E , y sea $L_k = \text{Span}_{\mathbb{R}} \{\vec{u}_1, \dots, \vec{u}_k\}$ el espacio vectorial generado por los k primeros vectores. Entonces, existe un conjunto de vectores $\vec{v}_1, \dots, \vec{v}_n$ tal que, para cada $k = 1, 2, \dots$, se verifica

1. El espacio vectorial $L'_k := \text{Span}_{\mathbb{R}} \{\vec{v}_1, \dots, \vec{v}_k\}$ coincide con L_k .
2. El vector $\vec{v}_{k+1}$ es ortogonal a L_k .

Observa que del teorema anterior se deduce que la recta de vector director $\vec{v}_{k+1}$ es ortogonal al espacio L_k (de dimensión k). Si denotamos por ℓ_{k+1} a esta recta, entonces escribimos abreviadamente $\ell_{k+1} \perp L_k$.

A continuación exponemos un algoritmo para calcular los vectores $\vec{v}_1, \dots, \vec{v}_n, \dots$ del teorema anterior. Es el llamado Algoritmo de ortogonalización de Gram-Schmidt.

2.1. Algoritmo de ortogonalización de Gram-Schmidt

El algoritmo de Gram-Schmidt consta de los siguientes pasos:

- **Input:** Un conjunto de vectores $B = \{\vec{u}_1, \dots, \vec{u}_n\}$ linealmente independiente.
- **Output:** Un conjunto de vectores $B_o = \{\vec{v}_1, \dots, \vec{v}_n\}$ linealmente independiente y ortogonales dos a dos.

1. Definimos $\vec{v}_1 := \vec{u}_1$.
2. Para $i = 2, \dots, n$, calculamos

$$\vec{v}_i = \vec{u}_i - \left(\frac{\langle \vec{u}_i, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 + \dots + \frac{\langle \vec{u}_i, \vec{v}_{i-1} \rangle}{\|\vec{v}_{i-1}\|^2} \vec{v}_{i-1} \right) \quad (18)$$

3. Definimos $B_o := \{\vec{v}_1, \dots, \vec{v}_n\}$.

Ejemplo 2.1. Sea $E = \mathbb{R}^2$, con el producto escalar definido por la fórmula

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (19)$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E . Supongamos que nos dan la base $B = \{\vec{u}_1 = (1, -1), \vec{u}_2 = (0, 1)\}$. Observa que

$$\langle \vec{u}_1, \vec{u}_2 \rangle = (1, -1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = (1, -1) \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -1 \neq 0.$$

Además $\|\vec{u}_1\|^2 = \langle \vec{u}_1, \vec{u}_1 \rangle = 1$. Aplicando el algoritmo de Gram-Schmidt a la base B obtenemos

$$\vec{v}_1 = \vec{u}_1 = (1, -1), \quad \vec{v}_2 = \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 = (0, 1) - \frac{-1}{1} (1, -1) = (1, 0).$$

Luego la base ortogonal asociada a B es

$$B_o = \{\vec{v}_1 = (1, -1), \vec{v}_2 = (1, 0)\}$$

En consecuenca la base ortonormal asociada a B según el algoritmo de Gram-Schmidt es

$$\tilde{B} = \left\{ \vec{e}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1 = (1, -1), \vec{e}_2 = \frac{1}{\|\vec{v}_2\|} \vec{v}_2 = (1, 0) \right\},$$

ya que $\|\vec{v}_2\| = 1$ para este productor escalar.

2.2. Proyecciones ortogonales

El concepto clásico de proyección de un vector de $\mathbb{R}^3$ sobre un plano que pasa por el origen admite una generación a espacios euclídeos de dimensión finita. También se puede generalizar en algunos casos a espacios de funciones de dimensión infinita.

Describiremos a continuación el concepto de proyección ortogonal sobre un subespacio W de un espacio euclídeo E .

Consideramos un espacio euclídeo E de dimensión n , y un subespacio W de dimensión $d < n$ sobre el que queremos proyectar. Se tiene que

$$E = W + W^\perp, \text{ con } W \cap W^\perp = \{\vec{0}\}. \quad (20)$$

Consideremos una base ortonormal $B_1 = \{\vec{e}_1, \dots, \vec{e}_d\}$ de W , respectivamente $B_2 = \{\vec{e}_{d+1}, \dots, \vec{e}_n\}$ base ortonormal de $W^\perp$. Entonces $B = B_1 \cup B_2$ es una base de E , y en esta base se tiene que las coordenadas de $\vec{u} \in E$ son

$$(\langle \vec{u}, \vec{e}_1 \rangle, \dots, \langle \vec{u}, \vec{e}_d \rangle, \langle \vec{u}, \vec{e}_{d+1} \rangle, \dots, \langle \vec{u}, \vec{e}_n \rangle). \quad (21)$$

por (16). Entonces la *proyección sobre W* queda definida por:

$$pr_W : E \rightarrow W, \quad pr_W(\vec{u}) := \langle \vec{u}, \vec{e}_1 \rangle \vec{e}_1 + \dots + \langle \vec{u}, \vec{e}_d \rangle \vec{e}_d. \quad (22)$$

Ilustraremos la construcción anterior sobre un ejemplo.

Ejemplo 2.2. Consideramos $\mathbb{R}^3$ con el producto escalar habitual. Sea W el plano engendrado por los vectores $\vec{v}_1 = (1, 0, -1)$, $\vec{v}_2 = (0, 1, -1)$. Una base ortonormal de W es

$$B_1 = \left\{ \vec{e}_1 = \frac{1}{\sqrt{2}}(1, 0, -1), \vec{e}_2 = \frac{1}{\sqrt{6}}(-1, 2, -1) \right\}$$

El complemento ortogonal de W es la recta $W^\perp$ engendada por $\vec{e}_3 = \frac{1}{\sqrt{3}}(1, 1, 1)$. En consecuencia, en la base $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ se tiene

$$pr_W : \mathbb{R}^3 \rightarrow W, \quad pr_W(\vec{u}) := \langle \vec{u}, \vec{e}_1 \rangle \vec{e}_1 + \langle \vec{u}, \vec{e}_2 \rangle \vec{e}_2.$$

Por ejemplo, si $\vec{u} = (1, 2, 3) = \langle \vec{u}, \vec{e}_1 \rangle \vec{e}_1 + \langle \vec{u}, \vec{e}_2 \rangle \vec{e}_2 + \langle \vec{u}, \vec{e}_3 \rangle \vec{e}_3 = \frac{-2}{\sqrt{2}}\vec{e}_1 + 0 \cdot \vec{e}_2 + \frac{6}{\sqrt{3}}\vec{e}_3$, se tiene que

$$pr_W(\vec{u}) = \frac{-2}{\sqrt{2}}\vec{e}_1 + 0 \cdot \vec{e}_2 = \frac{-2}{\sqrt{2}}\vec{e}_1 = (-1, 0, 1).$$

3. Aplicaciones ortogonales.

Consideraremos en esta sección un espacio euclídeo E con un productor escalar $\langle, \rangle$.

Definición 5. Sea $f : E \rightarrow E$ una aplicación lineal. Diremos que f es una aplicación ortogonal si satisface la siguiente condición:

$$\langle f(\vec{u}), f(\vec{v}) \rangle = \langle \vec{u}, \vec{v} \rangle, \quad \text{para cualesquiera } \vec{u}, \vec{v} \in E.$$

Nota 3.1. Se suele decir que una aplicación es ortogonal si conserva el producto escalar. Observa que en este caso conserva la longitud de los vectores, pues $\|\vec{u}\|^2 = \langle \vec{u}, \vec{u} \rangle = \langle f(\vec{u}), f(\vec{u}) \rangle = \|f(\vec{u})\|^2$.

En el caso de que el espacio euclídeo es $E = \mathbb{R}^n$ con el producto escalar usual, las aplicaciones ortogonales coinciden con las *matrices ortogonales*: matrices A de tamaño $n \times n$ tales que $AA^t = I_n$. El conjunto de matrices ortogonales se denota por $O(n; \mathbb{R})$,

$$O(n; \mathbb{R}) = \{A \in M_{n,n}(\mathbb{R}) \mid AA^t = I_n\}.$$

Ejemplo 3.1. Consideremos en $\mathbb{R}^3$ una base ortonormal $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, las siguientes transformaciones son ortogonales:

- La identidad, cuya matriz es $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- La simetría respecto del plano π de ecuación $x = 0$, cuya matriz es $S_\pi = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
- La simetría con respecto al eje $\ell = OX$, cuya matriz es $S_\ell = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.
- La simetría respecto al origen O (o simetría central), cuya matriz es $S_O = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

- El giro de eje OX y ángulo α , cuya matriz es $G_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$.
- El giro de eje OX y ángulo α compuesto con una simetría respecto del plano $x = 0$ (ò simetría rotatoria), cuya matriz es $\sigma_\alpha = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$. Observa que se da la igualdad $\sigma_\alpha = S_\ell \circ G_\alpha$.

Una propiedad **muy importante** de las matrices ortogonales es que las columnas (o filas) de una matriz ortogonal forman una base ortonormal (no solamente ortogonal).

4. Matrices unitarias.

Las matrices unitarias son el análogo de las matrices ortogonales en los complejos en lugar de en los reales. Concretamente, una matriz U de números complejos es unitaria si

$$UU^\dagger = U^\dagger U = I_n,$$

donde $U^\dagger$ es la matriz traspuesta **y conjugada**. Es decir, la propiedad de ser unitaria es como la de ser ortogonal si además se añade la conjugación.

Las matrices unitarias son muy útiles, por ejemplo, porque también *preservan un producto escalar complejo* (como en el caso de las matrices ortogonales arriba), que es el objeto matemático que se usa para expresar probabilidades en mecánica cuántica.

GUÍA DE ESTUDIO: ECUACIONES DIFERENCIALES MATRICIALES DE PRIMER ORDEN

Prof. Roberto Giménez Conejero

February 12, 2025

1. Ecuación diferencial lineal de orden n

Una *ecuación diferencial* es una ecuación matemática que relaciona una función con sus derivadas. El ejemplo más sencillo, es cuando una función es igual a su derivada,

$$f' = f.$$

La ecuación anterior tiene infinitas soluciones, $f(x) = Ce^x$, donde C es cualquier constante. Cuando damos una *condición inicial*, por norma general, estamos concretando una *solución particular* de la *solución general* anterior. Por ejemplo, si a la ecuación diferencial anterior le añadimos la condición de que $f(1) = 48$, entonces solo hay una solución, $f(x) = 48/e^x$, que obtenemos de la solución general anterior y añadiendo la condición:

$$Ce^1 = 48.$$

Una *ecuación diferencial lineal de orden n* es una ecuación diferencial donde aparece una función y sus n primeras derivadas (aunque alguna puede no aparecer, lo importante es que la n -ésima derivada sí aparece) y además aparecen de forma lineal, es decir, no hay potencias ni otro tipo de funciones que operen a las derivadas. De forma general, tienen esta estructura:

$$a_n f^{(n)} + a_{n-1} f^{(n-1)} + \dots + a_2 f^{(2)} + a_1 f^{(1)} + a_0 f = b,$$

donde los $f^{(i)}$ indican la i -ésima derivada de f , los a_i y b son funciones (al igual que f) que suponemos son conocidas.

Ejemplo 1.1. *La ecuación*

$$f^{(3)} - 3f = 49$$

es una ecuación diferencial lineal de orden 3, y la ecuación

$$\sin(x)f^{(42)} - \log(x)f^{(3)} - f = \tan(x)$$

es una ecuación diferencial lineal de orden 42. Sin embargo, la ecuación

$$(f^{(3)})^2 - 3\sin(f) = 49$$

es una ecuación diferencial no lineal.

Las ecuaciones diferenciales lineales se llaman así porque si f y g son soluciones de la ecuación diferencial, entonces $f + g$ y λf también son soluciones de la ecuación diferencial **cuando el término b es cero**, es decir, si la ecuación es *homogénea*. Esto es muy fácil de observar; si la ecuación original es

$$a_n f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_2 f^{(2)} + a_1 f^{(1)} + a_0 f = 0,$$

entonces para la suma de soluciones $h = f + g$ se tiene que

$$\begin{aligned} a_n (f + g)^{(n)} + a_{n-1} (f + g)^{(n-1)} + \cdots + a_2 (f + g)^{(2)} + a_1 (f + g)^{(1)} + a_0 (f + g) &= \\ a_n (f^{(n)} + g^{(n)}) + a_{n-1} (f^{(n-1)} + g^{(n-1)}) + \cdots + a_2 (f^{(2)} + g^{(2)}) + a_1 (f^{(1)} + g^{(1)}) + a_0 (f + g) &= \\ (a_n f^{(n)} + \cdots + a_1 f^{(1)} + a_0 f) + (a_n g^{(n)} + \cdots + a_1 g^{(1)} + a_0 g) &= 0 + 0, \end{aligned}$$

y de forma similar con λf , por las propiedades lineales de la derivada.

Hay varios métodos para resolver este tipo de ecuaciones directamente, que se pueden encontrar fácilmente en cualquier referencia sobre este tema, pero no vamos a explicarlos. Nos vamos a centrar en un método que transforma este tipo de ecuaciones en otro: en ecuaciones diferenciales matriciales de orden 1.

La idea es muy sencilla. Empezamos con f , que lo llamamos también y_0 (por 0-ésima derivada). Como la segunda derivada f'' es la derivada de f' , llamaremos $y_1 = f'$ y entonces se tiene que $f'' = y_1'$. De la misma forma, si denotamos f'' por y_2 , se tiene que $f''' = y_2'$. En general, tomamos la notación $y_i = f^{(i)}$ y lo que se tendrá es que $y_i' = y_{i+1}$. Esta idea nos permite cambiar la ecuación diferencial

$$a_n f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_2 f^{(2)} + a_1 f^{(1)} + a_0 f = b$$

por la ecuación

$$a_n y_n + a_{n-1} y_{n-1} + \cdots + a_2 y_2 + a_1 y_1 + a_0 y_0 = b,$$

si además añadimos las ecuaciones

$$\begin{aligned} y_1 &= y_0' \\ y_2 &= y_1' \\ &\vdots \\ y_{n-1} &= y_{n-2}' \\ y_n &= y_{n-1}'. \end{aligned}$$

La ventaja es que hemos cambiado derivadas de orden n a solo derivadas (de orden uno), pero ahora tenemos un sistema de ecuaciones diferenciales. En la siguiente sección vamos a estudiar cómo resolver este tipo de sistemas de ecuaciones diferenciales.

Nota. En realidad, en la siguiente sección vamos a explicar cómo resolver estos sistemas en el caso en el que $b = 0$ y $a_n = 1$. Además, el truco anterior para transformar una ecuación diferencial de orden n en un sistema de orden uno habría que cambiarlo ligeramente: habría que expresarlo todo en términos de $y_0, \dots, y_{n-1}$, para tener y_{n-1}' en la primera ecuación.

2. Ecuación diferencial matricial de orden 1

Como es habitual, en esta sección vamos a considerar x, y, z o, más en general, $y_1, \dots, y_n$, como funciones que dependen de una variable t (que normalmente omitiremos). Esto viene inspirado

en situaciones físicas, donde queremos estudiar una función que depende del tiempo.

Una ecuación diferencial matricial de orden 1 es el siguiente sistema de ecuaciones diferenciales:

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

donde, a priori, los a_{ij} son también funciones de t , aunque nosotros estudiaremos el caso en el que son constantes. En forma abreviada, esto lo escribiremos matricialmente como $Y' = AY$, donde Y es el vector columna con los y_i y A es la matriz de entradas a_{ij} . Además, igual que en el caso que se mencionaba anteriormente, normalmente habrá infinitas soluciones que se concretan si se dan *condiciones iniciales* en un tiempo t_0 :

$$\begin{pmatrix} y_1(t_0) \\ y_2(t_0) \\ \vdots \\ y_n(t_0) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n,$$

o, en notación matricial, $Y(t_0) = C$.

En el caso de que existan dos (vectores de) soluciones V_1, V_2 , es decir, que $V_i' = AV_i$, entonces también $\lambda_1 V_1 + \lambda_2 V_2$ será solución, por la linealidad de la derivada. Supongamos ahora que el tamaño de la matriz cuadrada y de los vectores columna es n , como escribíamos antes, y que además tenemos n (vectores de) soluciones $V_1, \dots, V_n$. En este caso, podemos obtener una matriz al juntar los vectores en columnas:

$$\begin{pmatrix} \vdots & \vdots & \cdots & \vdots \\ V_1 & V_2 & \cdots & V_n \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix}.$$

Ahora llegamos al resultado principal para encontrar soluciones de la ecuación diferencial matricial:

Teorema 1. *El determinante de la matriz anterior se llama Wronskiano, y es no nulo si, y solo si, los vectores $V_1, \dots, V_n$ son linealmente independientes. En ese caso, al conjunto $\{V_1, \dots, V_n\}$ lo llamamos conjunto fundamental de soluciones y cualquier solución de la ecuación diferencial matricial es una combinación lineal de estos, es decir, cualquier solución V tiene la forma*

$$V = b_1 V_1 + \cdots + b_n V_n,$$

para algunos valores $b_1, \dots, b_n$.

2.1. Buscar soluciones

En la práctica, buscar un conjunto fundamental de soluciones de la ecuación diferencial matricial $Y' = AY$ es muy sencillo, basta con seguir estos pasos:

1. Como la matriz A es una matriz de constantes, se pueden calcular los autovalores de la matriz y también los autovectores asociados.

2. Si un autovalor λ es real y tiene multiplicidad uno y autovector asociado $\vec{v}$, entonces la solución asociada a ese autovalor de la ecuación diferencial matricial es $V = e^{\lambda t} \vec{v}$.
3. El caso en el que los autovalores son complejos o su multiplicidad es mayor que uno no lo vamos a estudiar, pero es fácil encontrar cómo encontrar soluciones en ese caso.
4. El conjunto fundamental de soluciones se consigue tomando todas las soluciones obtenidas anteriormente.

Ejemplo 2.1. Para encontrar soluciones de la ecuación

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ 4 & -7 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

lo primero que hacemos es encontrar los autovalores asociados a la matriz. Para ello, calculamos las soluciones de

$$\begin{aligned} \det \left(\begin{pmatrix} 3 & -4 \\ 4 & -7 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) &= 0 \\ \det \begin{pmatrix} 3 - \lambda & -4 \\ 4 & -7 - \lambda \end{pmatrix} &= 0 \\ -21 - 3\lambda + 7\lambda + \lambda^2 + 16 &= 0. \end{aligned}$$

Las soluciones de la ecuación anterior son los autovalores: $\lambda_1 = 1$ y $\lambda_2 = -5$.

Ahora sacamos autovectores asociados. Para el autovalor 1, buscamos un vector $\vec{v}$ que satisfaga la ecuación del tipo $A\vec{v} = \lambda_1 \vec{v}$ o, equivalentemente, $(A - \lambda_1 I)\vec{v} = 0$:

$$\begin{pmatrix} 3 & -4 \\ 4 & -7 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 1 \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

con lo que después de resolver el sistema asociado obtenemos $\alpha = 2\beta$. Con esto, obtenemos el autovector

$$\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Haciendo el mismo tipo de cálculos para λ_2 , obtenemos el autovector

$$\vec{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Con estos datos, podemos obtener un conjunto fundamental de soluciones. Asociado al autovalor λ_1 es

$$V_1 = e^{\lambda_1 t} \vec{v}_1 = e^{1t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2e^t \\ e^t \end{pmatrix}.$$

De igual forma, para $\lambda_2 = -5$ es

$$V_2 = e^{\lambda_2 t} \vec{v}_2 = e^{-5t} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} e^{-5t} \\ 2e^{-5t} \end{pmatrix}.$$

Con esto, cualquier solución de la ecuación tiene la forma

$$V = b_1 V_1 + b_2 V_2 = b_1 \begin{pmatrix} 2e^t \\ e^t \end{pmatrix} + b_2 \begin{pmatrix} e^{-5t} \\ 2e^{-5t} \end{pmatrix}$$

para algunos valores $b_1, b_2 \in \mathbb{R}$. De hecho, podemos comprobar que esto es cierto con el teorema anterior, porque el Wronskiano es no nulo:

$$W = \det \begin{pmatrix} 2e^t & e^{t-5} \\ e^t & 2e^{-5t} \end{pmatrix} = 4e^{-4t} - e^{-4t} = 3e^{-4t},$$

que es no nulo.

Por último, solo queda estudiar qué ocurre cuando damos condiciones iniciales. Puesto que ya sabemos la forma de **cualquier solución**, solo tenemos que buscar una que cumpla las condiciones impuestas. Es decir, sabemos que una solución cualquiera tiene la forma

$$V = b_1 V_1 + b_2 V_2 + \cdots + b_n V_n$$

para algunos b_i , si además sabemos que nuestra solución debe cumplir que $V(t_0) = C$, esto nos determina los coeficientes por medio de un sistema de ecuaciones. Veámoslo con el ejemplo anterior.

Ejemplo 2.2. Volvamos al ejemplo anterior, con la misma ecuación y añadamos que la solución tiene como condiciones iniciales $x(0) = 47$ y que $y(0) = 13$. Matricialmente:

$$V(0) = \begin{pmatrix} 47 \\ 13 \end{pmatrix}.$$

Anteriormente ya habíamos calculado la solución genérica de la ecuación,

$$V = b_1 \begin{pmatrix} 2e^t \\ e^t \end{pmatrix} + b_2 \begin{pmatrix} e^{-5t} \\ 2e^{-5t} \end{pmatrix},$$

con lo que solo falta añadir nuestra condición:

$$V(0) = b_1 \begin{pmatrix} 2e^0 \\ e^0 \end{pmatrix} + b_2 \begin{pmatrix} e^{-5 \cdot 0} \\ 2e^{-5 \cdot 0} \end{pmatrix} = b_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + b_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 47 \\ 13 \end{pmatrix}.$$

Esto, claramente, se corresponde al sistema

$$\begin{cases} 2b_1 + b_2 = 47 \\ b_1 + 2b_2 = 13 \end{cases}$$

Si lo resolvemos, obtenemos que $b_1 = 27$ y $b_2 = -7$, con lo cual nuestra solución particular es

$$V(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = 27 \begin{pmatrix} 2e^t \\ e^t \end{pmatrix} - 7 \begin{pmatrix} e^{-5t} \\ 2e^{-5t} \end{pmatrix} = \begin{pmatrix} 54e^t - 7e^{-5t} \\ 27e^t - 14e^{-5t} \end{pmatrix}.$$

Además, podemos comprobarlo si quisiéramos volviendo a la ecuación diferencial inicial y usando que $x(t) = 54e^t - 7e^{-5t}$ y que $y(t) = 27e^t - 14e^{-5t}$.

Guía de estudio: Funciones de varias variables reales. Cálculo diferencial

Prof. Roberto Giménez Conejero

4 de marzo de 2025

En este documento presentamos una guía de estudio para el tema de Cálculo diferencial. Se puede ampliar su desarrollo en los temas correspondientes en los libros [3, 4, 1] y también [2].

1. Funciones en $\mathbb{R}^n$

Las funciones de varias variables surgen de manera natural en contextos muy elementales. Por ejemplo:

- $f(x, y) = \sqrt{x^2 + y^2}$ da la distancia del origen al punto (x, y) .
- $f(x, y) = xy$ da el área del rectángulo de dimensiones x e y .
- $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ da la distancia del origen al punto (x, y, z) .
- La magnitud de la fuerza gravitatoria ejercida por un cuerpo de masa M situado en el origen sobre un cuerpo de masa m situado en el punto (x, y, z) viene dada por $F(x, y, z) = \frac{GmM}{x^2+y^2+z^2}$.

Si el *dominio* de una función g no viene dado explícitamente, se sobreentiende que el dominio es el conjunto de todos los puntos para los cuales la definición de la función tiene sentido. La *imagen* o *rango* de g son todos aquellos valores que puede tomar g . Usaremos las siguientes notaciones para estos conceptos: si $g : U \rightarrow \mathbb{R}$ es

una función tal que $U \subset \mathbb{R}^n$, diremos que g es una función de n variables, que U es su dominio, su imagen o rango es

$$Im(g) = \{y \in \mathbb{R} \mid \text{existe un punto } p \in U \text{ tal que } y = g(p)\}.$$

Por ejemplo, si $g(x, y, z) = \arcsin(x + y + z)$, se entiende que el dominio es el conjunto de todos los puntos (x, y, z) para los cuales $-1 \leq x + y + z \leq 1$. Su imagen o rango es $Im(g) = [-\frac{\pi}{2}, \frac{\pi}{2}]$. La descripción de su dominio es la siguiente:

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid -1 - x - y \leq z \leq 1 - x - y\}.$$

1.1. Curvas de nivel. Superficies de nivel

La representación gráfica de funciones de varias variables es bastante distinta a la de una variable real. Estudiaremos los casos de dos y tres variables.

En el caso de funciones de dos variables ($n = 2$), si consideramos una función $f : U \rightarrow \mathbb{R}$ dibujaremos en primer lugar el mapa de la función. Este mapa $2D$ está formado por las curvas de nivel constante de la función f . Así si λ es el nivel fijado, la curva de nivel λ de f es

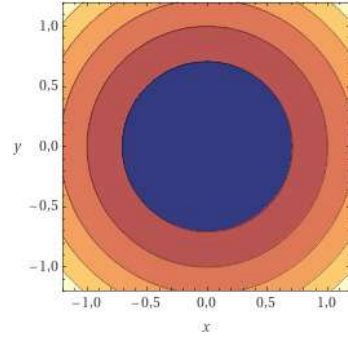
$$\mathcal{C}_\lambda = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = \lambda\} \quad (1)$$

La *gráfica o grafo de la función* es el subconjunto de $\mathbb{R}^3$ definido por

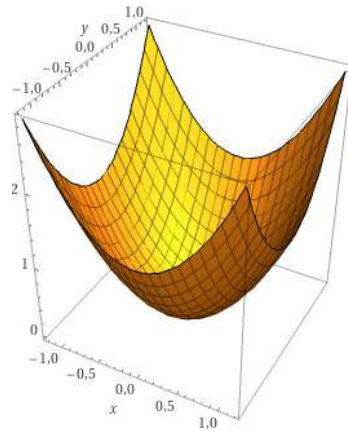
$$graf(f) = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\} = \bigcup_{\lambda \in \mathbb{R}} \mathcal{C}_\lambda \times \{\lambda\}. \quad (2)$$

Es decir, la gráfica se construye a partir del mapa poniendo cada curva de nivel a su nivel, esto es $\mathcal{C}_\lambda$ a altura $z = \lambda$.

Ejemplo 1.1. Sea $f(x, y) = x^2 + y^2$. Sus curvas de nivel son circunferencias de centro $(0, 0)$ y radio $\sqrt{\lambda}$, si $\lambda > 0$; la curva de nivel 0 es un tipo de curva degenerada ya que $\mathcal{C}_0 = \{(0, 0)\}$; finalmente, si $\lambda < 0$ no hay ningún punto real en la curva, es decir es el conjunto vacío. Su mapa tiene la representación gráfica siguiente:



Su gráfica es la siguiente figura:



Por otro lado, para funciones de tres variables ($n = 3$), si consideramos una función $f : U \rightarrow \mathbb{R}$ dibujaremos sólo el mapa de la función. Este mapa es ahora $3D$ y está formado por las superficies de nivel constante de la función f . Así si λ es el nivel fijado, la *superficie de nivel* λ de f es

$$\mathcal{S}_\lambda = \{(x, y, z) \in \mathbb{R}^3 \mid f(x, y, z) = \lambda\}. \quad (3)$$

Es este caso no podemos representar su gráfica ya que tendríamos que dibujarla en un espacio de cuatro dimensiones.

Ejemplo 1.2. Sea $f(x, y, z) = x^2 + y^2 + z^2$. Sus superficies de nivel son esferas de centro $(0, 0, 0)$ y radio $\sqrt{\lambda}$, si $\lambda > 0$; la superficie de nivel 0 es un tipo de superficie degenerada ya que $\mathcal{S}_0 = \{(0, 0, 0)\}$; finalmente, si $\lambda < 0$, no hay ningún punto real en la superficie $\mathcal{S}_\lambda$, es decir es el conjunto vacío.

En general, para funciones de n variables, si consideramos una función $f : U \rightarrow \mathbb{R}$ el mapa de la función es n dimensional y está formado por las *hipersuperficies de nivel constante de la función f* . Así si λ es el nivel fijado, la *hipersuperficie de nivel λ de f* es

$$\mathcal{H}_\lambda = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid f(x_1, \dots, x_n) = \lambda\} \quad (4)$$

La *gráfica o grafo de la función* es el subconjunto de $\mathbb{R}^{n+1}$ definido por

$$\text{graf}(f) = \{(x_1, \dots, x_n, z) \in \mathbb{R}^{n+1} \mid z = f(x_1, \dots, x_n)\} = \bigcup_{\lambda \in \mathbb{R}} \mathcal{H}_\lambda \times \{\lambda\}. \quad (5)$$

1.2. Límites y continuidad de funciones de varias variables reales

Presentaremos a continuación las nociones de límite y continuidad para funciones de 2 variables reales. Todos estos conceptos tienen sus análogos para funciones de $n \geq 2$ variables reales pero su exposición es menos intuitiva. Por otro lado, usaremos en estos espacios $\mathbb{R}^n$ la distancia euclídea usual dada por

$$d(P, Q) = \sqrt{(q_1 - p_1)^2 + \dots + (q_n - p_n)^2}, \quad (6)$$

donde $P = (p_1, \dots, p_n)$ y $Q = (q_1, \dots, q_n)$.

En primer lugar, para estudiar las funciones en varias variables se necesita considerar algunos tipos de conjuntos especiales. Los definiremos a continuación.

Definición 1.1. Sea $r > 0$ un número real positivo y $P \in \mathbb{R}^2$, definimos:

- La bola abierta de centro P y radio r , también llamada disco abierto, como:

$$B(P; r) := \{(x, y) \in \mathbb{R}^2 \mid d((x, y); P) < r\} \quad (7)$$

- La bola cerrada de centro P y radio r , también llamada disco cerrado, como:

$$\overline{B(P; r)} := \{(x, y) \in \mathbb{R}^2 \mid d((x, y); P) \leq r\} \quad (8)$$

- La bola abierta punteada de centro P y radio r , como:

$$B^*(P; r) := \{(x, y) \in \mathbb{R}^2 \mid 0 < ((x, y); P) < r\} \quad (9)$$

Consideremos una función de dos variables reales $f(x, y)$ que está definida alrededor de un punto $P = (a, b)$. Entonces, decimos que el *límite de f es L cuando (x, y) tiende a P* , y escribimos

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

si, para todo $\epsilon > 0$, existe $\delta > 0$ talque, para todo punto $(x, y) \in B^*(P; \delta)$, entonces $|f(x, y) - L| < \epsilon$.

Ejemplo 1.3. 1. La función dada por

$$f(x, y) = \frac{x^2 y}{x^2 + y^2} \quad (10)$$

tiene límite 0 en $(0, 0)$. Este hecho se deduce de la siguiente acotación:

$$0 \leq \left| \frac{x^2 y}{x^2 + y^2} \right| = \left| \frac{r^3 \cos^2 \theta \sin \theta}{r^2} \right| \leq r,$$

donde hemos usado el cambio a coordenadas polares en $\mathbb{R}^2$:

$$x = r \cos \theta, \quad y = r \sin \theta.$$

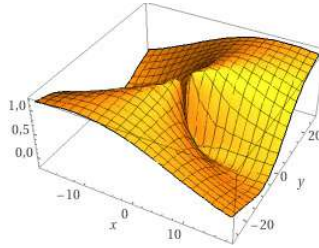
2. La función dada por

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

no tiene límite en $(0, 0)$. En efecto, sobre la recta de pendiente m la función toma el valor

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

valor que es distinto para cada recta $y = mx$, luego no existe este límite. Su gráfica es la siguiente figura:



Los límites de las funciones de varias variables verifican las siguientes propiedades, que sólo enunciaremos aquí para funciones de dos variables, pero que son aplicables a funciones de $n \geq 2$ variables.

Teorema 1.1. (*Propiedades de los límites*) Supongamos que $\lim_{(x,y) \rightarrow P} f(x, y)$ y $\lim_{(x,y) \rightarrow P} g(x, y)$ existen. Entonces:

1. *Propiedad de la suma:*

$$\lim_{(x,y) \rightarrow P} (f(x, y) + g(x, y)) = \lim_{(x,y) \rightarrow P} f(x, y) + \lim_{(x,y) \rightarrow P} g(x, y)$$

2. *Propiedad del múltiplo constante:* para cada constante $k \in \mathbb{R}$ se tiene

$$\lim_{(x,y) \rightarrow P} (k \cdot f(x, y)) = k \cdot \lim_{(x,y) \rightarrow P} f(x, y)$$

3. *Propiedad del producto:*

$$\lim_{(x,y) \rightarrow P} (f(x, y) \cdot g(x, y)) = \left(\lim_{(x,y) \rightarrow P} f(x, y) \right) \cdot \left(\lim_{(x,y) \rightarrow P} g(x, y) \right)$$

4. *Propiedad del cociente:* si $\lim_{(x,y) \rightarrow P} g(x, y) \neq 0$, entonces

$$\lim_{(x,y) \rightarrow P} \left(\frac{f(x, y)}{g(x, y)} \right) = \frac{\lim_{(x,y) \rightarrow P} f(x, y)}{\lim_{(x,y) \rightarrow P} g(x, y)}$$

La continuidad de funciones de varias variables es un poco más delicada que las de una variable real debido al tipo de entornos que tienen los puntos en $\mathbb{R}^n$, véase [1.1](#).

Definición 1.2. Una función $f(x, y)$ es continua en un punto (a, b) si

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

Se dice que es continua, si es continua en cada punto P de su dominio de definición.

Una propiedad útil de las funciones continuas es la siguiente:

Proposición 1.1. La composición de funciones continuas es continua.

2. Derivada direccional de una función. Aplicaciones

Consideremos una función $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^2$, y un vector unitario $\vec{u} = (u_1, u_2)$. En cada punto $P = (a, b)$ del dominio de f podemos considerar la variación de la función f en la dirección del vector $\vec{u}$. Matemáticamente esta variación se calcula como el siguiente límite:

$$D_{\vec{u}}f(P) := \lim_{h \rightarrow 0} \frac{f(a + hu_1, b + hu_2) - f(a, b)}{h}, \quad (11)$$

y se llama la derivada direccional de f en P en la dirección $\vec{u}$. Hay dos casos particulares importantes que tienen nombre propio, son las derivadas parciales de f , y que son casos particulares de la definición anterior, para $\vec{u} = \vec{e}_1 = (1, 0)$, y $\vec{u} = \vec{e}_2 = (0, 1)$.

Definición 2.1. Para una función de 2 variables $f : S \rightarrow \mathbb{R}$, definimos:

1. La derivada parcial de f con respecto a x como

$$\frac{\partial f}{\partial x}(P) = D_{\vec{e}_1}f(P) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h}. \quad (12)$$

2. La derivada parcial de f con respecto a y como

$$\frac{\partial f}{\partial y}(P) = D_{\vec{e}_2}f(P) = \lim_{h \rightarrow 0} \frac{f(a, b + h) - f(a, b)}{h}. \quad (13)$$

En este tema estudiaremos cómo calcular tales derivadas para una función f dada. También su importancia y uso para aproximar f cerca de P . Para esto usaremos el vector gradiente de f en P , que definimos a continuación.

Definición 2.2. Sean $F : S \rightarrow \mathbb{R}$ una función de varias variables reales con $S \subset \mathbb{R}^n$, y un punto $P \in S$. Entonces, **el gradiente de F** es el vector que tiene por componentes las derivadas parciales de F , es decir

$$\nabla F(P) := \left(\frac{\partial F}{\partial x_1}(P), \dots, \frac{\partial F}{\partial x_n}(P) \right). \quad (14)$$

Nota 2.1. En el caso de dos variables, la definición anterior es como sigue. Dada una función de 2 variables $f : S \rightarrow \mathbb{R}$, y un punto $P \in S$, definimos el gradiente de f en P , y lo denotamos por $\nabla f(P)$, como

$$\nabla f(P) := \left(\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P) \right). \quad (15)$$

En una siatición general, la derivada direccional de una función se calcula usando la fórmula (16) establecida en el siguiente teorema.

Teorema 2.1. (Cálculo de la derivada direccional) Sea F un función con derivadas parciales continuas en $\mathbb{R}^n$. Si $\vec{v}$ es un vector no nulo de $\mathbb{R}^n$, entonces $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$ es el vector unitario en la dirección de $\vec{v}$, y la derivada direccional de F en esta dirección viene dada por la fórmula:

$$D_{\vec{u}}F = \nabla F \cdot \vec{u} = \frac{\partial F}{\partial x_1}u_1 + \dots + \frac{\partial F}{\partial x_n}u_n \quad (16)$$

donde $\vec{u} = (u_1, \dots, u_n)$.

Ejemplo 2.1. Consideremos $F(x, y) = xe^y$, y $\vec{v} = (2, 3)$. Entonces, la derivada direccional de F en la dirección de $\vec{v}$ es

$$D_{\vec{u}}F = \frac{1}{\sqrt{2^2 + 3^2}} (e^y \cdot 2 + xe^y \cdot 3) = \frac{2 + 3x}{\sqrt{13}} e^y.$$

Nota 2.2. En el caso de funciones de dos variables $f : S \rightarrow \mathbb{R}^2$, el gradiente define una función (vectorial) $\nabla f : S \rightarrow \mathbb{R}^2$ dada por

$$(x, y) \longrightarrow \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right),$$

llamada habitualmente campo gradiente de la función f . En la siguiente sección analizaremos algunas propiedades de esta función.

2.1. El gradiente de una función

Sea $f : S \rightarrow \mathbb{R}$ una función de variable real, donde S es un abierto de $\mathbb{R}^n$. Supongamos que f tiene derivadas parciales $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq n$, continuas y consideramos el vector gradiente de f en $\mathbb{R}^n$ dado por

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

Es de señalar que si $\vec{u} = (u_1, \dots, u_n)$ es un vector unitario en $\mathbb{R}^n$, entonces la derivada direccional de f en la dirección $\vec{u}$ viene dada por la fórmula (16) y la del producto escala, es decir

$$D_{\vec{u}}f = \nabla f \cdot \vec{u} = \|\nabla f\| \cos(\angle(\nabla f, \vec{u})), \quad (17)$$

donde $\angle(\nabla f, \vec{u})$ denota el ángulo entre ∇f y $\vec{u}$. En consecuencia, tenemos las siguientes desigualdades que acotan los posibles valores de la derivada direccional de una función:

$$-\|\nabla f\| \leq D_{\vec{u}}f \leq \|\nabla f\|. \quad (18)$$

Dicho de otro modo, el vector gradiente apunta en la dirección y sentido de la máxima tasa de variación, y esta tasa máxima es $\|\nabla f\|$. Análogamente, la función f decrece más rápidamente en el sentido opuesto al gradiente de f , es decir, en la dirección $-\nabla f$. La tasa de variación de máximo decrecimiento de f es $-\|\nabla f\|$. Por otro lado, la derivada direccional es 0 en las direcciones ortogonales al gradiente. Para afianzar estas ideas mostramos cómo usarlas en unos ejemplos.

Ejemplo 2.2. Sea $f(x, y) = x^4 y^{-2}$ una función definida en $\mathbb{R}^2 \setminus \{y = 0\}$, y $P = (2, 1)$. Halla el vector unitario que apunte en la dirección y sentido de la máxima tasa de crecimiento de f en P .

Para calcular dicha tasa calcularemos el gradiente de la función y lo evaluaremos en P . Después tomaremos el vector unitario $\vec{u}$ correspondiente. Los cálculos son los siguientes.

$$\nabla f = (4x^3 y^{-2}, -2x^4 y^{-3}), \quad \nabla f(P) = (32, -32), \quad \vec{u} = (1/\sqrt{2}, 1/\sqrt{2}).$$

Ejemplo 2.3. La altitud de una montaña en variables (x, y) viene dada por la función

$$h(x, y) = 2500 + 100(x + y^2)e^{-0.3y^2},$$

donde x, y vienen dadas en unidades de 100 metros (m).

1. Halla la derivada direccional de f en el punto $P = (-1, -1)$ en la dirección del vector unitario $\vec{u}$ que forma un ángulo $\theta = \frac{\pi}{4}$ con el gradiente.

2. ¿Cuál es la interpretación de la derivada anterior?

La respuesta al primer apartado se halla usando la fórmula (17). En efecto,

$$\nabla f(-1, -1) = (74, -148), \quad D_{\vec{u}}f(-1, -1) = \|(74, -148)\| \cos\left(\frac{\pi}{4}\right) \simeq 165,5/\sqrt{2} \simeq 117.$$

La interpretación de esta derivada es sencilla. Si nos situamos en la montaña en el punto $(-1, -1, 2500)$, y empezamos a movernos sobre la montaña de forma que nuestro desplazamiento horizontal se realice en la dirección $\vec{u}$, entonces la altitud aumenta a razón de 117 metros por cada 100 metros de desplazamiento horizontal, o también 1,17 m por cada metro.

Otra propiedad importante del campo gradiente es que es perpendicular a las curvas de nivel de la función. Este hecho es una consecuencia de la regla de la cadena que estudiaremos a continuación, véase Proposición 2.1. Resumimos lo expuesto en esta sección en el siguiente teorema.

Teorema 2.2. (Interpretación geométrica del gradiente) Sea $f : S \rightarrow \mathbb{R}$ una función de variable real, donde S es un abierto de $\mathbb{R}^2$. Supongamos que f tiene derivadas parciales $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq 2$, continuas y que el vector gradiente de f en el punto $P \in S$ no es el vector nulo. Sea $\vec{u}$ un vector unitario formando un ángulo θ con $\nabla f(P)$. Entonces

$$D_{\vec{u}}f(P) = \|\nabla f(P)\| \cos(\theta).$$

Además, se verifica

- $\nabla f(P)$ apunta en la dirección y sentido de la tasa de variación máxima de crecimiento de f en P .
- $-\nabla f(P)$ apunta en la dirección y sentido de la tasa de variación máxima de decrecimiento de f en P .
- $\nabla f(P)$ es ortogonal a la curva de nivel de f en P .

Nota 2.3. Es un ejercicio interesante la formulación y verificación del resultado análogo al anterior para funciones $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^n$, que se propone al lector.

2.2. Regla de la cadena en varias variables

Recordemos en primer lugar que si una función de una variable, $y = f(x)$, y la variable x varía con el tiempo t , es decir $x = x(t)$, entonces la fórmula de derivación llamada “Regla de la cadena” proporciona el cálculo de la variación de y como función de t . En concreto:

$$\frac{dy}{dt} = \frac{df}{dx} = \frac{df}{dx} \cdot \frac{dx}{dt}.$$

Por ejemplo, si $y = \cos(3x^2+1)$, y la variable x es la función dada por $x = x(t) = 2t^3$, entonces $\dot{y} = \frac{dy}{dt} = -\sin(12t^6+1) \cdot (6t^2)$. Comenzaremos estudiando la generalización de esta fórmula en el caso de varias variables. Es la llamada “Regla de la cadena para trayectoria”.

Proposición 2.1. (*Regla de la cadena para una trayectoria*) Sea $F : S \rightarrow \mathbb{R}$ una función de varias variables reales con $S \subset \mathbb{R}^n$, y un punto $P \in S$. Supongamos dada una curva en S con punto inicial P , es decir una función $\vec{x}(t)$ derivable y con derivada continua definida en un intervalo $[a, b] \subset \mathbb{R}$ tal que $\vec{x}(a) = P$,

$$[a, b] \rightarrow S, \quad \vec{x}(t) = (x_1(t), \dots, x_n(t)). \quad (19)$$

Entonces, la derivada respecto de la variable t de la función compuesta $F \circ \vec{x}(t)$ se calcula mediante la fórmula:

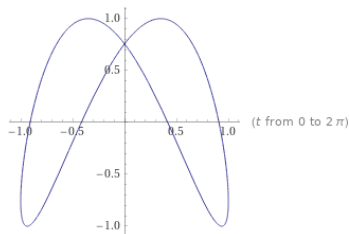
$$\frac{dF}{dt}(a) = \nabla F(P) \cdot \frac{d\vec{x}}{dt}(a) = \frac{\partial F}{\partial x_1}(P) \frac{dx_1}{dt}(a) + \dots + \frac{\partial F}{\partial x_n}(P) \frac{dx_n}{dt}(a). \quad (20)$$

La proposición anterior en el caso de dos variables (x, y) admite la siguiente formulación:

$$\frac{dF}{dt} = \nabla F \cdot \frac{d\vec{x}}{dt} = \frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt}. \quad (21)$$

Ilustraremos su aplicación sobre un ejemplo.

Ejemplo 2.4. Consideremos la función $F(x, y) = 20 + 10e^{-0.3(x^2+y^2)}$ que modeliza la función temperatura sobre una placa (en $^{\circ}\text{C}$), y la trayectoria $\vec{c} = (\cos(t-2), \sin(2t))$ (t en segundos), cuya gráfica se puede ver en la siguiente figura:



Entonces

$$\nabla F = \left(-6xe^{-0,3(x^2+y^2)}, -6ye^{-0,3(x^2+y^2)} \right), \quad \frac{d\vec{c}}{dt} = (-\sin(t-2), 2\cos(2t)) .$$

En consecuencia,

$$\frac{dF}{dt} = 6x \sin(t-2)e^{-0,3(x^2+y^2)} - 12y \cos(2t)e^{-0,3(x^2+y^2)} .$$

Si nos interesara la tasa de variación de la función temperatura F en el punto $P = \vec{c}(0) = (\cos(2), 0) \simeq (0,9, 0)$ al desplazarnos sobre la curva $\vec{c}(t)$ obtendríamos $\frac{dF}{dt}(0) \simeq 2,155^\circ\text{C/s}$.

Proposición 2.2. Sea $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^2$ abierto y f derivadas parciales continuas. Entonces, para cada punto $P \in S$, el vector $\nabla f(P)$ es perpendicular a la curva de nivel de f que pasa por P .

Demostración. En efecto, si $\vec{c}(t) = (x(t), y(t))$ es una parametrización de la curva de nivel de f que pasa por P , entonces, $f \circ \vec{c}(t) = \lambda$ para una cierta constante λ ; luego, por (21), se tiene que

$$\frac{df \circ \vec{c}}{dt} = \frac{d\lambda}{dt} = 0 \implies \nabla f(P) \cdot \vec{c}'(t) = 0,$$

lo que da el resultado, ya que $\vec{c}'(t)$ es el vector tangente a la curva de nivel λ . $\square$

La versión general de la Regla de la cadena (no sólo para una trayectoria) establece el siguiente resultado.

Teorema 2.3. Regla de la cadena para funciones. Sea $f : S \rightarrow \mathbb{R}$ una función definida en un conjunto abierto $S \subset \mathbb{R}^n$. Supongamos que cada función coordenada x_i es a su vez una función diferenciable de un conjunto de m variables reales, $\underline{t} = (t_1, \dots, t_m)$, para $i = 1, \dots, n$. Si denotamos por $\vec{x}(\underline{t}) = (x_1(\underline{t}), \dots, x_n(\underline{t}))$, entonces tenemos la siguiente fórmula de derivación:

$$\frac{\partial f \circ \vec{x}}{\partial t_k}(\underline{t}) = \nabla f \cdot \left(\frac{\partial x_1}{\partial t_k}, \dots, \frac{\partial x_n}{\partial t_k} \right) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t_k} \quad (22)$$

Ilustraremos con un ejemplo cómo se aplica la fórmula (22).

Ejemplo 2.5. Consideremos la función $f(x, y) = x^2 + y^2 + x$, y el cambio de coordenadas a polares dado por

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Se pide calcular la variación de f con respecto a la variable radio r y también respecto a la variable angular θ . Para esto aplicaremos la fórmula (22) para las variables

$$(x_1, x_2) = (x, y), \quad (t_1, t_2) = (r, \theta).$$

En consecuencia

$$\frac{\partial f}{\partial r} = (2x + 1, 2y) \cdot (\cos \theta, \sin \theta) = (2r \cos \theta + 1, 2r \sin \theta) \cdot (\cos \theta, \sin \theta) = 2r + \cos \theta,$$

$$\frac{\partial f}{\partial \theta} = (2x + 1, 2y) \cdot (-r \sin \theta, r \cos \theta) = (2r \cos \theta + 1, 2r \sin \theta) \cdot (-r \sin \theta, r \cos \theta) = -r \sin \theta.$$

2.2.1. Derivación de funciones definidas de forma implícita

Supondamos dada una superficie Γ en $\mathbb{R}^3$ definida por la ecuación $F(x, y, z) = 0$. Supongamos que cerca de un punto P de Γ la coordenada z se puede poner como función de (x, y) , es decir $z = z(x, y)$, en este caso, $\frac{\partial F}{\partial x} \neq 0$, y se dan las siguiente fórmulas, llamadas de *Fórmulas de derivación en implícitas*

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \quad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}, \quad (23)$$

donde hemos usado las notaciones $F_x = \frac{\partial F}{\partial x}$, $F_y = \frac{\partial F}{\partial y}$, $F_z = \frac{\partial F}{\partial z}$. La verificación de la fórmula anterior es fácil: basta aplicar la regla de la cadena, (22), a la ecuación $F(x, y, z) = 0$. Mostraremos su uso en el siguiente ejemplo.

Ejemplo 2.6. Si consideramos el hiperboloide de una hoja de ecuación

$$F(x, y, z) = x^2 + y^2 - 2z^2 + 12x - 8z - 4 = 0,$$

entonces las variaciones $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$ en el punto $P = (1, 1, 1)$, vienen dadas por

$$\frac{\partial z}{\partial x}(P) = -\frac{2x + 12}{-4z - 8} \Bigg|_P = \frac{7}{6}, \quad \frac{\partial z}{\partial y}(P) = -\frac{2y}{-4z - 8} \Bigg|_P = \frac{1}{6}.$$

2.3. Diferenciabilidad y continuidad de funciones en varias variables

Consideremos una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Diremos que f es **diferenciable en un punto P_0 de $\mathbb{R}^n$** , si y sólo si existe una matriz T de tamaño $m \times n$ tal que

$$\lim_{X \rightarrow P_0} \frac{\|f(X) - [f(P_0) + T(\vec{P_0 X})]\|}{\|\vec{P_0 X}\|} = 0. \quad (24)$$

Llamaremos a la matriz T la derivada o diferencial de f en P_0 , y la denotaremos por $D_{P_0}f$, o también $Df(P_0)$. Se da el siguiente resultado.

Teorema 2.4. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en un punto $P_0 \in \mathbb{R}^n$. Entonces todas las derivadas parciales de f existen en P_0 y se tiene que la matriz $T = (t_{ij})$ de tamaño $m \times n$ que satisface la condición (24) es*

$$T = (t_{ij}) = \left(\frac{\partial f_i}{\partial x_j}(P_0) \right),$$

donde $f = (f_1, \dots, f_m)$ son las componentes de f y $(x_1, \dots, x_n)$ es un sistema de coordenadas en $\mathbb{R}^n$.

Demostración. Véase [3], pag. 158. □

En el caso de funciones de varias variables, también se da el hecho de que la diferenciabilidad de f en un punto implica la continuidad en ese punto. En concreto, usando el siguiente resultado se puede demostrar el teorema 2.5.

Lema 2.1. Véase [3], pag 162.

Teorema 2.5. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en un punto $P_0 \in \mathbb{R}^n$. Entonces f es continua en P_0 , y además existe una constante $M > 0$ tal que*

$$\|f(X) - f(P_0)\| < M\|X - P_0\|, \quad (25)$$

para X cerca de P_0 y $X \neq P_0$.

Demostración. Véase [3], pag. 158. □

Finalmente podemos concluir que, en el caso de que la función considerada f tenga derivadas parciales continuas, entonces f es diferenciable. En concreto, para el caso de dos variables tendríamos el siguiente enunciado.

Teorema 2.6. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable. Entonces todas las derivadas parciales existen, y la derivada direccional en un punto X en la dirección $\vec{u}$ está dada por la fórmula:*

$$D_{\vec{u}}f(X) = \nabla f(X) \cdot \vec{u} \quad (26)$$

Demostración. Consideremos la curva $\vec{c}(t) = X + t\vec{u}$. □

2.4. Aproximación de funciones y Teorema de Taylor

2.4.1. Plano tangente a una superficie en un punto. Aproximación lineal

Consideremos dada una superficie en $\mathbb{R}^3$ definida por una ecuación $F(x, y, z) = 0$, es decir $\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 \mid F(x, y, z) = 0\}$. Consideremos un punto $P_0 = (x_0, y_0, z_0)$ de dicha superficie, es decir $F(x_0, y_0, z_0) = 0$. Entonces, se define *el plano tangente a $\mathcal{S}$ en P_0* como el plano definido por la ecuación:

$$\frac{\partial F}{\partial x}(P_0)(x - x_0) + \frac{\partial F}{\partial y}(P_0)(y - y_0) + \frac{\partial F}{\partial z}(P_0)(z - z_0) = 0 . \quad (27)$$

Un caso particular importante es el de considerar las superficies que son gráficas de funciones de dos variables, $f(x, y)$. En este caso $F(x, y, z) = f(x, y) - z$, y la ecuación anterior tiene una fórmula más sencilla:

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - (z - z_0) = 0 , \quad (28)$$

donde $z_0 = f(x_0, y_0)$.

Ejemplo 2.7. 1. Si $F(x, y, z) = x^2 + y^2 + z^2 - 1$, entonces el plano tangente a la superficie $F = 0$ en el punto $P_0 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ es el plano

$$\frac{2}{\sqrt{2}}(x - \frac{1}{\sqrt{2}}) + \frac{2}{\sqrt{2}}(y - \frac{1}{\sqrt{2}}) = 0 . \quad (29)$$

2. Si $f(x, y) = x^2 - y^2$ y consideramos la superficie que da su gráfica, entonces, por ejemplo, el plano tangente a la gráfica en el punto (de la gráfica $P_0 = (1, 1, 0)$) es

$$2(x - 1) - 2(y - 1) - z = 0 . \quad (30)$$

Una utilidad de la ecuación del plano tangente a una gráfica de la forma $z = f(x, y)$ es para construir la aproximación lineal de la función f alrededor del punto (x_0, y_0) . Se tiene entonces la siguiente fórmula:

$$f(x, y) \simeq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0), \quad (31)$$

fórmula que recibe el nombre de *aproximación lineal de f alrededor del punto (x_0, y_0)* .

Ejemplo 2.8. Si $f(x, y) = x^2 - y^2$, y consideramos el punto $P = (1, 1)$, la aproximación lineal de f alrededor del punto P es

$$f(x, y) \simeq 0 + 2(x - 1) - 2(y - 1) = 2x - 2y. \quad (32)$$

2.4.2. Derivadas de orden superior. Teorema de Taylor

Las derivadas parciales de orden superior son las derivadas parciales de las derivadas parciales de una función f . Así, si $f : S \rightarrow \mathbb{R}$ es una función de dos variables, por ejemplo (x, y) , las derivadas parciales de segundo orden son las derivadas parciales de $f_x = \frac{\partial f}{\partial x}$ y de $f_y = \frac{\partial f}{\partial y}$; es decir,

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

En las fórmulas anteriores se aprecia la importancia del orden de derivación iterada a la hora de calcular las *derivadas cruzadas* f_{yx} , f_{xy} . Sin embargo, para funciones muy generales tal ambigüedad desaparece debido al , también conocido como teorema de Schwarz.

Teorema 2.7. Si f_{yx} y f_{xy} son continuas en una bola B , entonces $f_{yx} = f_{xy}$ en cada punto de B .

Ilustraremos el resultado anterior sobre un ejemplo.

Ejemplo 2.9. Consideremos $f(x, y) = e^{x/y}$. Entonces $f_x = \frac{1}{y}e^{x/y}$, $f_y = -\frac{x}{y^2}e^{x/y}$. Entonces las derivadas segundas de f son:

$$f_{xx} = \frac{1}{y^2}e^{x/y}, \quad f_{yx} = \left(-\frac{1}{y^2} - \frac{x}{y^3}\right)e^{x/y},$$

$$f_{yy} = \left(\frac{2x}{y^3} + \frac{x}{y^4}\right)e^{x/y}, \quad f_{xy} = \left(-\frac{x}{y^3} - \frac{1}{y^2}\right)e^{x/y},$$

lo que comprueba la igualdad $f_{yx} = f_{xy}$.

Las derivadas de orden superior son un elemento importante de las funciones para construir sus aproximaciones polinómicas. En concreto, esto sucede con el teorema de Taylor 2.8 para funciones en varias variables.

La **notación** que usaremos es la siguiente, llamada notación de multiíndice. Si $\alpha = (\alpha_1, \dots, \alpha_n)$ es un vector de números naturales de tamaño n y $\mathbf{v} = (v_1, \dots, v_n)$, tomamos la siguiente notación:

$$\begin{aligned} |\alpha| &= \alpha_1 + \dots + \alpha_n, \\ \alpha! &= \alpha_1! \dots \alpha_n!, \\ \mathbf{v}^\alpha &= v_1^{\alpha_1} \dots v_n^{\alpha_n}. \end{aligned}$$

Además, si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función cuyas derivadas parciales de cualquier orden son diferenciables (para que no haya problema con el orden de las parciales), tomamos la notación:

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha} = \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Con esta notación, el teorema de Taylor queda de la siguiente forma.

Teorema 2.8. (*Teorema de Taylor para funciones de varias variables*)

Sea $f : S \rightarrow \mathbb{R}$ una función. Supongamos que f y todas sus derivadas parciales de cualquier orden diferenciables en un punto $\mathbf{a} \in \mathbb{R}^n$ (por el orden de las parciales). Entonces, si $\mathbf{x} = (x_1, \dots, x_n)$, existen funciones $E_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ para cada α con $|\alpha| = k$ tal que

$$\begin{aligned} f(\mathbf{x}) &= P_{k,\mathbf{a}}(\mathbf{x}) + \sum_{|\alpha|=k} E_\alpha(\mathbf{x})(\mathbf{x} - \mathbf{a})^\alpha \\ &= \sum_{|\alpha| \leq k} \frac{D^\alpha f(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha + \sum_{|\alpha|=k} E_\alpha(\mathbf{x})(\mathbf{x} - \mathbf{a})^\alpha, \end{aligned}$$

donde además

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} E_\alpha(\mathbf{x}) = 0$$

para todos los E_α .

En el teorema anterior, $P_{k,\mathbf{a}}(\mathbf{x})$ se llama **el polinomio de Taylor de orden k de f en el punto $\mathbf{a}$** , y $E_\alpha(\mathbf{x})$ son las **funciones error** (que son muy pequeños cuando $\mathbf{x}$ está muy cerca del punto $\mathbf{a}$).

Dada la complejidad de la fórmula anterior la explicitaremos para el caso particular de funciones de dos variables $n = 2$ y orden de diferenciabilidad $m = 3$. En este caso el teorema anterior se puede reformular como sigue.

Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función. Supongamos que f y todas sus derivadas parciales son diferenciables en $\mathbf{a} = (x_0, y_0)$. Entonces, el polinomio de Taylor de orden 2 de f en $\mathbf{a}$ es

$$P_{2,\mathbf{a}}(x, y) = f(\mathbf{a}) + \frac{\partial f}{\partial x}(\mathbf{a})(x - x_0) + \frac{\partial f}{\partial y}(\mathbf{a})(y - y_0) + \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) \frac{(x - x_0)^2}{2!} + \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a})(x - x_0)(y - y_0) + \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \frac{(y - y_0)^2}{2!}$$

y el polinomio de Taylor de orden 3 de f en $\mathbf{a}$ es, escribiendo $v_1 = x - x_0$ y $v_2 = y - y_0$,

$$P_{3,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \frac{\partial f}{\partial x}(\mathbf{a})v_1 + \frac{\partial f}{\partial y}(\mathbf{a})v_2 + \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) \frac{v_1^2}{2!} + \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a})v_1v_2 + \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \frac{v_2^2}{2!} + \frac{\partial^3 f}{\partial x^3}(\mathbf{a}) \frac{v_1^3}{3!} + \frac{\partial^3 f}{\partial x^2 \partial y}(\mathbf{a}) \frac{v_1^2v_2}{2!} + \frac{\partial^3 f}{\partial x \partial y^2}(\mathbf{a}) \frac{v_1v_2^2}{2!} + \frac{\partial^3 f}{\partial y^3}(\mathbf{a}) \frac{v_2^3}{3!}.$$

La fórmula de P_2 proporciona **la aproximación cuadrática de la función f cerca del punto (x_0, y_0)** , y se puede reescribir como

$$P_{2,(x_0,y_0)}(x, y) = f(x_0, y_0) + \nabla f(\underline{a}) \cdot (x - x_0, y - y_0) + \frac{1}{2}(x - x_0, y - y_0) \begin{pmatrix} f_{xx}(\underline{a}) & f_{yx}(\underline{a}) \\ f_{xy}(\underline{a}) & f_{yy}(\underline{a}) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \quad (33)$$

La matriz 2×2 de la fórmula anterior se llama **la matriz Hessiana de f** y será fundamental para el estudio de los puntos extremos de la función f . Se denota por $Hess f$, y si la evaluamos en un punto $\underline{a}$, entonces escribiremos $Hess_{\underline{a}} f$.

Ejemplo 2.10. *Calcula la fórmula de aproximación cuadrática (33) para la función $f(x, y) = x^3 + y^3 - x$ en el punto $P = (1, -1)$.*

En primer lugar calculamos las derivadas necesarias:

$$\nabla f = (3x^2 - 1, 3y^2), \quad Hess f = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}.$$

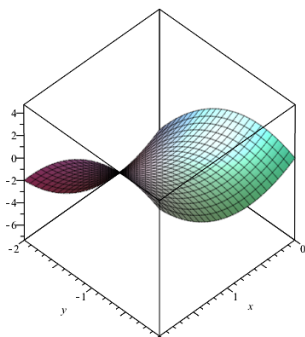
Después evaluamos en el punto dado:

$$f(P) = -1, \quad \nabla f(P) = (2, 3), \quad Hess_P f = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix},$$

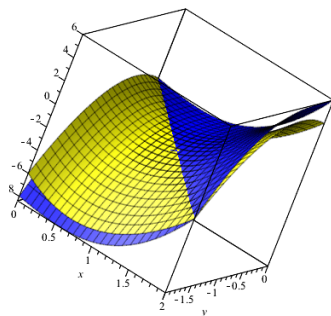
lo que da la aproximación cuadrática:

$$P_{2,P}(x, y) = -1 + 2(x-1) + 3(y+1) + \frac{1}{2}(x-1, y+1) \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} \begin{pmatrix} x-1 \\ y+1 \end{pmatrix} = 3x^2 - 3y^2 - 4x - 3y.$$

Observa que la superficie $z = 3x^2 - 3y^2 - 4x - 3y$ es un parabolide hiperbólico y tiene la representación gráfica siguiente



y proporciona una aproximación cuadrática local de la función f cerca del punto $P = (1, -1)$. Por ejemplo, usando el comando `plot3d([x^3+y^3-x, 3x^2-3y^2-4x-3y], x = 0 .. 2, y = -2 .. 0, color = [blue, yellow])` de Maple podemos representar ambas gráficas y se observa lo siguiente:



2.5. Máximos y mínimos de funciones de varias variables

La optimización es el proceso de hallar los valores extremos de una función. Esto consiste en hallar por ejemplo los puntos donde se alcanzan valores máximos y mínimos de la gráfica sobre un cierto dominio S . Análogamente al caso de una variable, es importante distinguir entre extremos locales y globales. Para tratar este tema necesitaremos establecer algunas definiciones básicas.

Definición 2.3. Una función $f : S \rightarrow \mathbb{R}$ definida en un dominio S de $\mathbb{R}^n$ tiene un extremo local en un punto $P \in S$ si existe una bola centrada en P , $B(P; \epsilon) \subset S$ tal que se da alguna de las dos situaciones siguientes

1. Para todo punto $\underline{x} \in B(P; \epsilon)$ se verifica la desigualdad $f(\underline{x}) \leq f(P)$. Situación llamada de **máximo local** en P .
2. Para todo punto $\underline{x} \in B(P; \epsilon)$ se verifica la desigualdad $f(\underline{x}) \geq f(P)$. Situación llamada de **mínimo local** en P .

Definición 2.4. Un punto $P \in S$ se dice un punto crítico de f si el gradiente de f en P es el vector nulo, es decir $\nabla f(P) = \vec{0}$, o alguna derivada parcial de f no existe en el punto P .

Los conceptos anteriores están relacionados. El siguiente teorema (atribuido a Fermat) establece su conexión.

Teorema 2.9. Si f tiene un máximo o un mínimo local en un punto P , entonces P es un punto crítico de f .

Por simplicidad haremos la demostración en el caso de funciones de dos variables, aunque el resultado es cierto para cualquier número n de variables.

Demostración. Supongamos en primer lugar que f tiene un mínimo local en $P = (x_0, y_0)$. Entonces en una cierta bola B centrada en P se tiene que $f(x, y) \geq f(P)$. Entonces la función $g(x) := f(x, y_0)$ tiene un mínimo en $x = x_0$, y por tanto $g'(x_0) = 0$, puesto que $g'(x) = \frac{\partial f}{\partial x}(x, y_0)$. Luego $\frac{\partial f}{\partial x}(x_0, y_0) = 0$. Además, si definimo la función auxiliar $h(y) := f(x_0, y)$, se tiene que $h'(y) = \frac{\partial f}{\partial y}(x_0, y)$, y entonces h tiene un mínimo en y_0 . Por esto $h'(y_0) = 0$, lo que implica $\frac{\partial f}{\partial y}(x_0, y_0) = 0$. En consecuencia $\nabla f(P) = \vec{0}$.

Un razonamiento análogo se puede hacer para demostrar el enunciado en el caso de que P sea un máximo local de f . Se propone al lector completar los detalles de la demostración en este caso. $\square$

Ilustraremos el método de cálculo de extremos locales mediante un ejemplo.

Ejemplo 2.11. Sea $f(x, y) = 11x^2 - 2xy + 2y^2 + 3y$. Calcula los puntos críticos de f .

En primer lugar calcularemos el gradiente de la función:

$$\nabla f = (22x - 2y, -2x + 4y + 3),$$

para imponer la condición $\nabla f = \vec{0}$. Lo que nos proporciona el sistema de ecuaciones

$$22x - 2y = 0, -2x + 4y + 3 = 0 ,$$

cuya única solución es $x = -1/14$, $y = -11/14$. Luego el extremo es el punto $P = (-1/14, -11/14)$. El valor de la función en este punto es $f(P) = -33/28$.

En el caso de funciones de dos variables se puede establecer un criterio análogo al que se tiene en el caso de las funciones de una variable para determinar el tipo de extremo de la función. Lo enunciaremos a continuación.

Teorema 2.10. (*Test de las segundas derivadas*)

Sea P un punto crítico de una función $f(x, y)$. Supongamos que existen las derivadas segundas f_{xx} , f_{yx} , f_{yy} y son continuas cerca de P . Consideremos la matriz Hessiana de f en P :

$$Hess_P f = \begin{pmatrix} f_{xx}(P) & f_{yx}(P) \\ f_{xy}(P) & f_{yy}(P) \end{pmatrix} .$$

Sea $D := \det(Hess_P f)$. Entonces:

1. Si $D > 0$, entonces
 - Si $f_{xx}(P) > 0$, el punto P es un mínimo local de f .
 - Si $f_{xx}(P) < 0$, el punto P es un máximo local de f .
2. Si $D < 0$, entonces P es un punto de silla.
3. Si $D = 0$, el criterio no decide.

Es de señalar que si aplicamos este criterio a la función del ejemplo 2.11 obtenemos que $P = (-1/14, -11/14)$ es un mínimo de la función ya que $D = 84 > 0$ y también $f_{xx}(P) = 22 > 0$. Quedaría por determinar si este extremo es global o sólomente local.

Teorema 2.11. Sea f una función continua sobre un dominio cerrado y acotado $\mathcal{D} \subset \mathbb{R}^2$. Entonces

- La función f alcanza un valor máximo y uno mínimo en $\mathcal{D}$.
- Los extremos ocurren o bien en los puntos críticos del interior de $\mathcal{D}$, o bien en los puntos de la frontera de $\mathcal{D}$.

Ejemplo 2.12. Halla el valor máximo de la función $f(x, y) = 2x + y - 3xy$ sobre el cuadrado unidad $\mathcal{D} = [0, 1] \times [0, 1]$.

1. Examina los puntos críticos.
2. Analiza los puntos en la frontera.
3. Compara.

Ejemplo 2.13. Halla el volumen máximo de una caja inscrita en el tetraedro limitado por los planos coordenados y el plano $\frac{1}{3}x + y + x = 1$.

1. Halla la función a maximizar.
2. Examina los puntos críticos.
3. Analiza los puntos en la frontera.
4. Compara.

2.5.1. Extremos condicionados de funciones

Teorema 2.12. Supongamos dadas dos funciones diferenciables $f(x, y)$, $g(x, y)$. Si f tiene un mínimo o un máximo local sobre la curva dada implícitamente por la ecuación $g(x, y) = 0$ en el punto $P = (x_0, y_0)$ tal que $\nabla g(P) \neq \vec{0}$, entonces existe un escalar λ tal que

$$\nabla f(P) = \lambda \nabla g(P) . \quad (34)$$

Ejemplo 2.14. Halla los valores críticos de la función $f(x, y) = 2x + 5y$ sobre la elipse de ecuación:

$$\left(\frac{x}{4}\right)^2 + \left(\frac{y}{3}\right)^2 = 1.$$

Procederemos de la manera siguiente:

1. Escribe la ecuación de Lagrange (34).
2. Aísla λ como función de (x, y) .
3. Halla x_0 e y_0 usando la restricción.

3. Aplicaciones

3.1. Cordenadas polares en $\mathbb{R}^2$. Coordenadas cilíndricas y esféricas en $\mathbb{R}^3$

3.2. Operadores diferenciales y cambios de coordenadas

3.2.1. Operadores diferenciales: Gradiente, Laplaciano, Rotacional

Referencias

- [1] T. M. Apostol. *Análisis Matemático*. Reverté, 1982.
- [2] E. Hernández, M. J. Vázquez, and M. A. Zurro. *Álgebra Lineal y Geometría*. Pearson, 2012.
- [3] J.E. Marsden and A.J. Tromba. *Cálculo Vectorial*. Addison-Wesley Iberoamericana, 2004.
- [4] J. Rogawski. *Cálculo. Varias Variables*. Reverté, 2012.

Guía de estudio: Funciones de varias variables reales.

Prof. Roberto Giménez Conejero

4 de marzo de 2025

Se puede ampliar este tema en los temas correspondientes en los libros [3, 4, 1] y también [2].

1. Espacios euclídeos

El contenido de esta sección sigue la línea de presentación dada en [2].

Un *espacio euclídeo* es un espacio vectorial real E dotado de un producto escalar, es decir, de una regla que asigne a cada par de vectores $\vec{u}, \vec{v}$ un número real designado por $\langle \vec{u}, \vec{v} \rangle$, de forma que se cumplen las siguientes propiedades:

1. Propiedad simétrica: $\langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle$.
2. Propiedad distributiva: $\langle \vec{u} + \vec{w}, \vec{v} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{w}, \vec{v} \rangle$.
3. $\langle \lambda \vec{u}, \vec{v} \rangle = \lambda \langle \vec{u}, \vec{v} \rangle$ para cada $\lambda \in \mathbb{R}$.
4. $\langle \vec{u}, \vec{v} \rangle \geq 0$, y sólo es 0 si $\vec{u} = \vec{v}$.

Se define el módulo o longitud (o norma) de un vector $\vec{u} \in E$ como

$$\|\vec{u}\| = \sqrt{\langle \vec{u}, \vec{u} \rangle}. \quad (1)$$

Ejemplo 1.1. 1. Si $E = \mathbb{R}^2$, y $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, entonces la fórmula

$$\langle \vec{u}, \vec{v} \rangle := x_1 x_2 + y_1 y_2 \quad (2)$$

define un producto escalar en E .

2. Si $E = \mathbb{R}^2$, y $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$, entonces la fórmula

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (3)$$

define un producto escalar en E . En este caso, el módulo de un vector tiene la fórmula

$$\|\vec{u}\| = \sqrt{x_1^2 + 2x_1y_1 + 2y_1^2} = \sqrt{(x_1 + y_1)^2 + y_1^2}.$$

3. Si E es el espacio de funciones continuas reales en el intervalo $[a, b]$, denotado por $C([a, b])$, se define para cada par de funciones f, g en E el producto escalar

$$\langle f, g \rangle := \int_a^b f(t)g(t) dt, \quad (4)$$

obteniéndose así un espacio euclídeo de funciones.

En un espacio euclídeo se cumple la siguiente desigualdad.

Proposición 1.1 (Desigualdad de Schwarz). *En un espacio euclídeo E , se da*

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \cdot \|\vec{v}\|,$$

para cualesquiera $\vec{u}, \vec{v}$ en E .

Demostración. Consideremos dos vectores cualesquiera $\vec{u}, \vec{v}$ en E , y λ un escalar real arbitrario. Entonces,

$$\langle \lambda \vec{u} - \vec{v}, \lambda \vec{u} - \vec{v} \rangle \geq 0, \quad (5)$$

por las propiedades del producto escalar. Desarrollando la desigualdad anterior se obtiene

$$\lambda^2 \|\vec{u}\|^2 - 2 \langle \vec{u}, \vec{v} \rangle + \|\vec{v}\|^2 \geq 0.$$

La parte izquierda de esta ecuación es un polinomio cuadrático en λ que no puede tener raíces reales y distintas. En consecuencia su discriminante ha de ser siempre ≤ 0 . Por tanto $\langle \vec{u}, \vec{v} \rangle^2 - \|\vec{u}\|^2 \|\vec{v}\|^2 \leq 0$. Tomando raíces cuadradas obtenemos $|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\| \|\vec{v}\|$, que es la desigualdad pedida. $\square$

Ortogonalidad. Sea E un espacio euclídeo con producto escalar $\langle, \rangle$. Dados dos vectores $\vec{u}, \vec{v}$ en E , diremos que son ortogonales (respecto de $\langle, \rangle$) si $\langle \vec{u}, \vec{v} \rangle = 0$.

Además si W es un subespacio de E , se define su complemento ortogonal $W^\perp$ como el conjunto

$$W^\perp = \{ \vec{w} \in E \mid \langle \vec{w}, \vec{u} \rangle = 0 \text{ para todo } \vec{u} \in W \}. \quad (6)$$

En estos espacios existe una noción de *ángulo* entre vectores. En concreto, si $\vec{u}, \vec{v}$ están en E , definimos el *coseno del ángulo* que forman $\vec{u}$ y $\vec{v}$ como

$$\cos \angle(\vec{u}, \vec{v}) := \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \cdot \|\vec{v}\|}. \quad (7)$$

En consecuencia, el *ángulo* que forman $\vec{u}$ y $\vec{v}$ es

$$\angle(\vec{u}, \vec{v}) := \arccos \left(\frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\| \cdot \|\vec{v}\|} \right) \in [0, \pi]. \quad (8)$$

Ejemplo 1.2. Si $E = \mathbb{R}^2$, y consideramos el producto escalar definido por la fórmula

$$\langle \vec{u}, \vec{v} \rangle := (x_1, y_1) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \quad (9)$$

para $\vec{u} = (x_1, y_1)$, $\vec{v} = (x_2, y_2)$ en E , entonces

$$\angle((1, 0), (0, 1)) = \arccos \left(\frac{1}{1 \cdot \sqrt{2}} \right) = \frac{\pi}{4}. \quad (10)$$

2. Funciones en $\mathbb{R}^n$

Las funciones de varias variables surgen de manera natural en contextos muy elementales. Por ejemplo:

- $f(x, y) = \sqrt{x^2 + y^2}$ da la distancia del origen al punto (x, y) .
- $f(x, y) = xy$ da el área del rectángulo de dimensiones x e y .
- $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ da la distancia del origen al punto (x, y, z) .
- La magnitud de la fuerza gravitatoria ejercida por un cuerpo de masa M situado en el origen sobre un cuerpo de masa m situado en el punto (x, y, z) viene dada por $F(x, y, z) = \frac{GmM}{x^2 + y^2 + z^2}$.

Si el *dominio* de una función g no viene dado explícitamente, se sobreentiende que el dominio es el conjunto de todos los puntos para los cuales la definición de la función tiene sentido. La *imagen* o *rango* de g son todos aquellos valores que puede tomar g . Usaremos las siguientes notaciones para estos conceptos: si $g : U \rightarrow \mathbb{R}$ es una función tal que $U \subset \mathbb{R}^n$, diremos que g es una función de n variables, que U es su dominio, su imagen o rango es

$$Im(g) = \{y \in \mathbb{R} \mid \text{existe un punto } p \in U \text{ tal que } y = g(p)\}.$$

Por ejemplo, si $g(x, y, z) = \arcsin(x + y + z)$, se entiende que el dominio es el conjunto de todos los puntos (x, y, z) para los cuales $-1 \leq x + y + z \leq 1$. Su imagen o rango es $Im(g) = [-\frac{\pi}{2}, \frac{\pi}{2}]$. La descripción de su dominio es la siguiente:

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid -1 - x - y \leq z \leq 1 - x - y\}.$$

2.1. Curvas de nivel. Superficies de nivel

La representación gráfica de funciones de varias variables es bastante distinta a la de una variable real. Estudiaremos los casos de dos y tres variables.

En el caso de funciones de dos variables ($n = 2$), si consideramos una función $f : U \rightarrow \mathbb{R}$ dibujaremos en primer lugar el mapa de la función. Este mapa $2D$ está formado por las curvas de nivel constante de la función f . Así si λ es el nivel fijado, la *curva de nivel* λ de f es

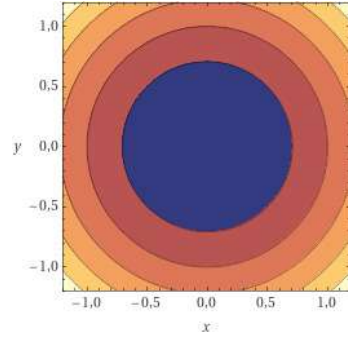
$$\mathcal{C}_\lambda = \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = \lambda\} \quad (11)$$

La *gráfica* o *grafo* de la función es el subconjunto de $\mathbb{R}^3$ definido por

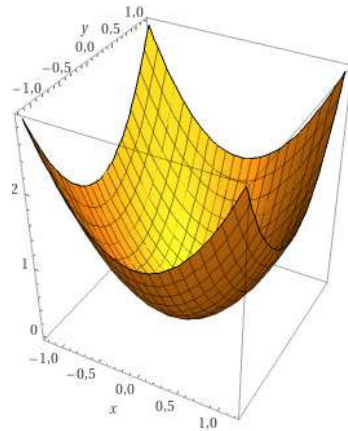
$$graf(f) = \{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\} = \bigcup_{\lambda \in \mathbb{R}} \mathcal{C}_\lambda \times \{\lambda\}. \quad (12)$$

Es decir, la gráfica se construye a partir del mapa poniendo cada curva de nivel a su nivel, esto es $\mathcal{C}_\lambda$ a altura $z = \lambda$.

Ejemplo 2.1. Sea $f(x, y) = x^2 + y^2$. Sus curvas de nivel son circunferencias de centro $(0, 0)$ y radio $\sqrt{\lambda}$, si $\lambda > 0$; la curva de nivel 0 es un tipo de curva degenerada ya que $\mathcal{C}_0 = \{(0, 0)\}$; finalmente, si $\lambda < 0$ no hay ningún punto real en la curva, es decir es el conjunto vacío. Su mapa tiene la representación gráfica siguiente:



Su gráfica es la siguiente figura:



Por otro lado, para funciones de tres variables ($n = 3$), si consideramos una función $f : U \rightarrow \mathbb{R}$ dibujaremos sólo el mapa de la función. Este mapa es ahora $3D$ y está formado por las superficies de nivel constante de la función f . Así si λ es el nivel fijado, la *superficie de nivel* λ de f es

$$\mathcal{S}_\lambda = \{(x, y, z) \in \mathbb{R}^3 \mid f(x, y, z) = \lambda\}. \quad (13)$$

Es este caso no podemos representar su gráfica ya que tendríamos que dibujarla en un espacio de cuatro dimensiones.

Ejemplo 2.2. Sea $f(x, y, z) = x^2 + y^2 + z^2$. Sus superficies de nivel son esferas de centro $(0, 0, 0)$ y radio $\sqrt{\lambda}$, si $\lambda > 0$; la superficie de nivel 0 es un tipo de superficie degenerada ya que $\mathcal{S}_0 = \{(0, 0, 0)\}$; finalmente, si $\lambda < 0$, no hay ningún punto real en la superficie $\mathcal{S}_\lambda$, es decir es el conjunto vacío.

En general, para funciones de n variables, si consideramos una función $f : U \rightarrow \mathbb{R}$ el mapa de la función es n dimensional y está formado por las *hipersuperficies de nivel constante de la función* f . Así si λ es el nivel fijado, la *hipersuperficie de nivel* λ de f es

$$\mathcal{H}_\lambda = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid f(x_1, \dots, x_n) = \lambda\} \quad (14)$$

La *gráfica o grafo de la función* es el subconjunto de $\mathbb{R}^{n+1}$ definido por

$$\text{graf}(f) = \{(x_1, \dots, x_n, z) \in \mathbb{R}^{n+1} \mid z = f(x_1, \dots, x_n)\} = \bigcup_{\lambda \in \mathbb{R}} \mathcal{H}_\lambda \times \{\lambda\}. \quad (15)$$

2.2. Límites y continuidad de funciones de varias variables reales

Presentaremos a continuación las nociones de límite y continuidad para funciones de 2 variables reales. Todos estos conceptos tienen sus análogos para funciones de $n \geq 2$ variables reales pero su exposición es menos intuitiva. Por otro lado, usaremos en estos espacios $\mathbb{R}^n$ la distancia euclídea usual dada por

$$d(P, Q) = \sqrt{(q_1 - p_1)^2 + \dots + (q_n - p_n)^2}, \quad (16)$$

donde $P = (p_1, \dots, p_n)$ y $Q = (q_1, \dots, q_n)$.

En primer lugar, para estudiar las funciones en varias variables se necesita considerar algunos tipos de conjuntos especiales. Los definiremos a continuación.

Definición 2.1. Sea $r > 0$ un número real positivo y $P \in \mathbb{R}^2$, definimos:

- La bola abierta de centro P y radio r , también llamada disco abierto, como:

$$B(P; r) := \{(x, y) \in \mathbb{R}^2 \mid d((x, y); P) < r\} \quad (17)$$

- La bola cerrada de centro P y radio r , también llamada disco cerrado, como:

$$\overline{B(P; r)} := \{(x, y) \in \mathbb{R}^2 \mid d((x, y); P) \leq r\} \quad (18)$$

- La bola abierta punteada de centro P y radio r , como:

$$B^*(P; r) := \{(x, y) \in \mathbb{R}^2 \mid 0 < ((x, y); P) < r\} \quad (19)$$

Consideremos una función de dos variables reales $f(x, y)$ que está definida alrededor de un punto $P = (a, b)$. Entonces, decimos que el *límite de f es L cuando (x, y) tiende a P* , y escribimos

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

si, para todo $\epsilon > 0$, existe $\delta > 0$ talque, para todo punto $(x, y) \in B^*(P; \delta)$, entonces $|f(x, y) - L| < \epsilon$.

Ejemplo 2.3. 1. La función dada por

$$f(x, y) = \frac{x^2 y}{x^2 + y^2} \quad (20)$$

tiene límite 0 en $(0, 0)$. Este hecho se deduce de la siguiente acotación:

$$0 \leq \left| \frac{x^2 y}{x^2 + y^2} \right| = \left| \frac{r^3 \cos^2 \theta \sin \theta}{r^2} \right| \leq r,$$

donde hemos usado el cambio a coordenadas polares en $\mathbb{R}^2$:

$$x = r \cos \theta, \quad y = r \sin \theta.$$

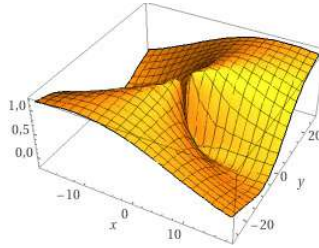
2. La función dada por

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

no tiene límite en $(0, 0)$. En efecto, sobre la recta de pendiente m la función toma el valor

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

valor que es distinto para cada recta $y = mx$, luego no existe este límite. Su gráfica es la siguiente figura:



Los límites de las funciones de varias variables verifican las siguientes propiedades, que sólo enunciaremos aquí para funciones de dos variables, pero que son aplicables a funciones de $n \geq 2$ variables.

Teorema 2.1. (*Propiedades de los límites*) Supongamos que $\lim_{(x,y) \rightarrow P} f(x, y)$ y $\lim_{(x,y) \rightarrow P} g(x, y)$ existen. Entonces:

1. *Propiedad de la suma:*

$$\lim_{(x,y) \rightarrow P} (f(x, y) + g(x, y)) = \lim_{(x,y) \rightarrow P} f(x, y) + \lim_{(x,y) \rightarrow P} g(x, y)$$

2. *Propiedad del múltiplo constante:* para cada constante $k \in \mathbb{R}$ se tiene

$$\lim_{(x,y) \rightarrow P} (k \cdot f(x, y)) = k \cdot \lim_{(x,y) \rightarrow P} f(x, y)$$

3. *Propiedad del producto:*

$$\lim_{(x,y) \rightarrow P} (f(x, y) \cdot g(x, y)) = \left(\lim_{(x,y) \rightarrow P} f(x, y) \right) \cdot \left(\lim_{(x,y) \rightarrow P} g(x, y) \right)$$

4. *Propiedad del cociente:* si $\lim_{(x,y) \rightarrow P} g(x, y) \neq 0$, entonces

$$\lim_{(x,y) \rightarrow P} \left(\frac{f(x, y)}{g(x, y)} \right) = \frac{\lim_{(x,y) \rightarrow P} f(x, y)}{\lim_{(x,y) \rightarrow P} g(x, y)}$$

La continuidad de funciones de varias variables es un poco más delicada que las de una variable real debido al tipo de entornos que tienen los puntos en $\mathbb{R}^n$, véase 2.1.

Definición 2.2. Una función $f(x, y)$ es continua en un punto (a, b) si

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

Se dice que es continua, si es continua en cada punto P de su dominio de definición.

Una propiedad útil de las funciones continuas es la siguiente:

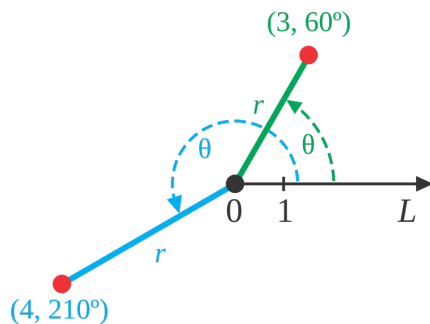
Proposición 2.1. La composición de funciones continuas es continua.

2.3. Otras coordenadas en $\mathbb{R}^2$ y $\mathbb{R}^3$

En ciertas situaciones, es conveniente cambiar de coordenadas para tener algo más fácil con lo que trabajar.

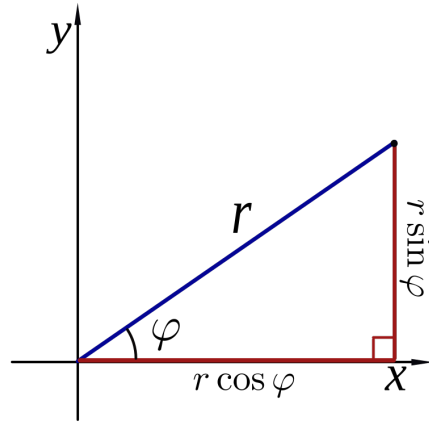
2.3.1. Coordenadas polares en $\mathbb{R}^2$

Cualquier punto en el plano $\mathbb{R}^2$ se puede determinar por sus coordenadas, por ejemplo (p, q) , o también por sus *coordenadas polares*, esto es, es suficiente dar la distancia al origen y el ángulo con el eje OX para saber dónde está el punto. Por ejemplo, los siguientes puntos:



están determinados de forma unívoca, como se ve en la imagen. Por otro lado, es importante saber cómo pasar información en coordenadas (cartesianas), digamos (x, y) , a coordenadas polares, digamos (r, θ) , y viceversa, lo cual es muy sencillo

usando las propiedades del seno y el coseno:



Es decir, el cambio entre unas coordenadas es el siguiente:

De polares (r, θ) a cartesianas (x, y) :

$$x = r \cos(\theta)$$

$$y = r \sin(\theta).$$

De cartesianas (x, y) a polares (r, θ) :

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \begin{cases} \arccos\left(\frac{x}{r}\right) & \text{si } y \geq 0 \text{ y } r \neq 0 \\ -\arccos\left(\frac{x}{r}\right) & \text{si } y < 0 \\ \text{no se define, se toma } 0 & \text{si } r = 0. \end{cases}$$

En la práctica, uno calcula $\arccos\left(\frac{x}{r}\right)$ y ve qué signo es el que tiene sentido para el ángulo.

Ejemplo 2.4. Para convertir $(4, \frac{2\pi}{3})$ en coordenadas cartesianas, basta con tomar

$$x = 4 \cos\left(\frac{2\pi}{3}\right) = -2$$

$$y = 4 \sin\left(\frac{2\pi}{3}\right) = 2\sqrt{3}.$$

Para convertir $(-1, -1)$ en coordenadas polares primero se calcula r

$$r = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}.$$

Y después θ . Primero calculamos el arcocoseno correspondiente,

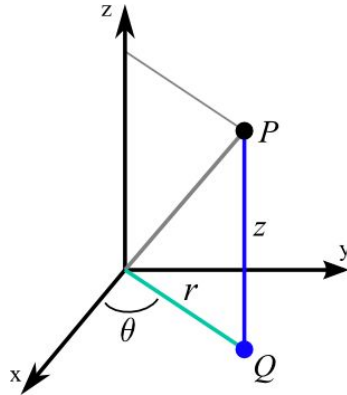
$$\arccos\left(\frac{-1}{\sqrt{2}}\right) = \frac{3\pi}{4},$$

que como es un ángulo que está en el segundo cuadrante pero nuestro punto sabemos que está en el tercer cuadrante, debe estar en negativo. Total:

$$\theta = -\frac{3\pi}{4}.$$

2.3.2. Coordenadas cilíndricas en $\mathbb{R}^3$

En $\mathbb{R}^3$ se puede determinar un punto si se conoce su coordenada z y, con las dos coordenadas (x, y) se usan coordenadas polares (puesto que determinar dos coordenadas se puede hacer con polares, como hemos visto antes). Esto se llama coordenadas cilíndricas en $\mathbb{R}^3$, que usan las coordenadas polares para determinar un punto en el plano XY de $\mathbb{R}^3$ y una coordenada adicional para determinar la componente z :



Cambiar entre coordenadas cilíndricas y cartesianas es muy sencillo si se conoce cómo cambiar a coordenadas polares en el plano.

De cilíndricas (r, θ, z) a cartesianas (x, y, z) :

$$x = r \cos(\theta)$$

$$y = r \sin(\theta)$$

$$z = z.$$

De cartesianas (x, y, z) a polares (r, θ, z) :

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \begin{cases} \arccos\left(\frac{x}{r}\right) & \text{si } y \geq 0 \text{ y } r \neq 0 \\ -\arccos\left(\frac{x}{r}\right) & \text{si } y < 0 \\ \text{no se define, se toma } 0 & \text{si } r = 0. \end{cases}$$

$$z = z$$

Ejemplo 2.5. Para convertir $(4, \frac{2\pi}{3}, 431897)$ en coordenadas cartesianas, basta con tomar

$$x = 4 \cos\left(\frac{2\pi}{3}\right) = -2$$

$$y = 4 \sin\left(\frac{2\pi}{3}\right) = 2\sqrt{3}$$

$$z = 431897.$$

Para convertir $(-1, -1, 983)$ en coordenadas cilíndricas primero se calcula r

$$r = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}.$$

Y después θ . Primero calculamos el arcocoseno correspondiente,

$$\arccos\left(\frac{-1}{\sqrt{2}}\right) = \frac{3\pi}{4},$$

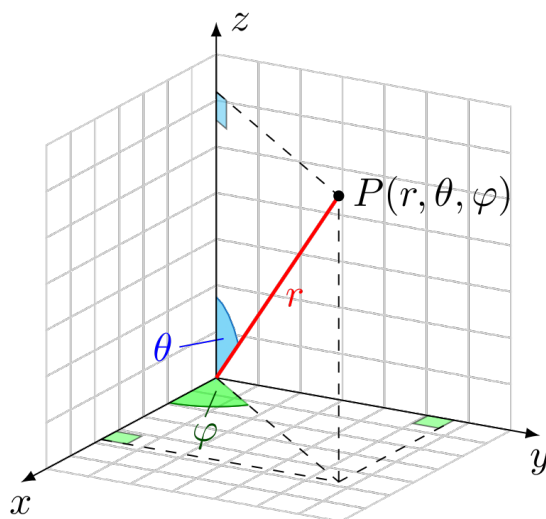
que como es un ángulo que está en el segundo cuadrante pero nuestro punto sabemos que está en el tercer cuadrante, debe estar en negativo. Total:

$$\theta = -\frac{3\pi}{4}.$$

Además $z = 983$.

2.3.3. Coordenadas esféricas en $\mathbb{R}^3$

Veamos primero la imagen de las coordenadas esféricas y después explicamos coordenada a coordenada:



- r es la distancia del punto al origen.
- φ es el ángulo que se calculaba también en coordenadas cilíndricas, es decir, proyectamos el punto en el plano XY (o, lo que es lo mismo, ignoramos la coordenada z) y calculamos el ángulo con el eje OX .
- θ es el ángulo que forma el punto (es decir, el vector que va entre el origen y el punto) y el eje OZ .

Estas tres cosas determinan el punto. Veamos cómo cambiar de coordenadas cartesianas a esféricas.

De esféricas (r, φ, θ) a cartesianas (x, y, z) :

$$x = r \cos(\varphi) \sin(\theta)$$

$$y = r \sin(\varphi) \sin(\theta)$$

$$z = r \cos(\theta).$$

De cartesianas (x, y, z) a esféricas (r, φ, θ) :

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2} \\ \varphi &= \text{signo}(y) \arccos \frac{x}{\sqrt{x^2 + y^2}} \\ \theta &= \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \end{aligned}$$

Ejemplo 2.6. Para convertir $(4, \frac{2\pi}{3}, \frac{\pi}{2})$ en coordenadas cartesianas, basta con tomar

$$\begin{aligned} x &= 4 \cos\left(\frac{2\pi}{3}\right) \sin\left(\frac{\pi}{2}\right) = -2 \\ y &= 4 \sin\left(\frac{2\pi}{3}\right) \sin\left(\frac{\pi}{2}\right) = 2\sqrt{3} \\ z &= 4 \cos\left(\frac{2\pi}{3}\right) = -2. \end{aligned}$$

Para convertir $(1, -1, \sqrt{2})$ en coordenadas esféricas primero se calcula r

$$r = \sqrt{1^2 + (-1)^2 + \sqrt{2}^2} = \sqrt{4} = 2.$$

Y después φ y θ :

$$\begin{aligned} \varphi &= \text{signo}(-1) \arccos \frac{1}{\sqrt{1^2 + (-1)^2}} = -\frac{\pi}{4} \\ \theta &= \arccos \frac{\sqrt{2}}{\sqrt{1^2 + (-1)^2 + \sqrt{2}^2}} = \frac{\pi}{4}. \end{aligned}$$

Referencias

- [1] T. M. Apostol. *Análisis Matemático*. Reverté, 1982.
- [2] E. Hernández, M. J. Vázquez, and M. A. Zurro. *Álgebra Lineal y Geometría*. Pearson, 2012.
- [3] J.E. Marsden and A.J. Tromba. *Cálculo Vectorial*. Addison-Wesley Iberoamericana, 2004.
- [4] J. Rogawski. *Cálculo. Varias Variables*. Reverté, 2012.

Guía de estudio: Cálculo diferencial de funciones de varias variables reales.

Prof. Roberto Giménez Conejero

18 de marzo de 2025

Se puede ampliar este tema en los temas correspondientes en los libros [3, 4, 1] y también [2].

1. Derivada direccional de una función. Aplicaciones

Consideremos una función $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^2$, y un vector unitario $\vec{u} = (u_1, u_2)$. En cada punto $P = (a, b)$ del dominio de f podemos considerar la variación de la función f en la dirección del vector $\vec{u}$. Matemáticamente esta variación se calcula como el siguiente límite:

$$D_{\vec{u}}f(P) := \lim_{h \rightarrow 0} \frac{f(a + hu_1, b + hu_2) - f(a, b)}{h}, \quad (1)$$

y se llama *la derivada direccional de f en P en la dirección $\vec{u}$* . Hay dos casos particulares importantes que tienen nombre propio, son *las derivadas parciales de f* , y que son casos particulares de la definición anterior, para $\vec{u} = \vec{e}_1 = (1, 0)$, y $\vec{u} = \vec{e}_2 = (0, 1)$.

Definición 1.1. Para una función de 2 variables $f : S \rightarrow \mathbb{R}$, definimos:

1. La derivada parcial de f con respecto a x como

$$\frac{\partial f}{\partial x}(P) = D_{\vec{e}_1}f(P) = \lim_{h \rightarrow 0} \frac{f(a + h, b) - f(a, b)}{h}. \quad (2)$$

2. La derivada parcial de f con respecto a y como

$$\frac{\partial f}{\partial y}(P) = D_{\vec{e}_2}f(P) = \lim_{h \rightarrow 0} \frac{f(a, b+h) - f(a, b)}{h}. \quad (3)$$

En este tema estudiaremos cómo calcular tales derivadas para una función f dada. También su importancia y uso para aproximar f cerca de P . Para esto usaremos el vector gradiente de f en P , que definimos a continuación.

Definición 1.2. Sean $F : S \rightarrow \mathbb{R}$ una función de varias variables reales con $S \subset \mathbb{R}^n$, y un punto $P \in S$. Entonces, **el gradiente de F** es el vector que tiene por componentes las derivadas parciales de F , es decir

$$\nabla F(P) := \left(\frac{\partial F}{\partial x_1}(P), \dots, \frac{\partial F}{\partial x_n}(P) \right). \quad (4)$$

Nota 1.1. En el caso de dos variables, la definición anterior es como sigue. Dada una función de 2 variables $f : S \rightarrow \mathbb{R}$, y un punto $P \in S$, definimos el gradiente de f en P , y lo denotamos por $\nabla f(P)$, como

$$\nabla f(P) := \left(\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P) \right). \quad (5)$$

En una siatición general, la derivada direccional de una función se calcula usando la fórmula (6) establecida en el siguiente teorema.

Teorema 1.1. (Cálculo de la derivada direccional) Sea F un función con derivadas parciales continuas en $\mathbb{R}^n$. Si $\vec{v}$ es un vector no nulo de $\mathbb{R}^n$, entonces $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$ es el vector unitario en la dirección de $\vec{v}$, y la derivada direccional de F en esta dirección viene dada por la fórmula:

$$D_{\vec{u}}F = \nabla F \cdot \vec{u} = \frac{\partial F}{\partial x_1}u_1 + \dots + \frac{\partial F}{\partial x_n}u_n \quad (6)$$

donde $\vec{u} = (u_1, \dots, u_n)$.

Ejemplo 1.1. Consideremos $F(x, y) = xe^y$, y $\vec{v} = (2, 3)$. Entonces, la derivada direccional de F en la dirección de $\vec{v}$ es

$$D_{\vec{u}}F = \frac{1}{\sqrt{2^2 + 3^2}} (e^y \cdot 2 + xe^y \cdot 3) = \frac{2 + 3x}{\sqrt{13}} e^y.$$

Nota 1.2. En el caso de funciones de dos variables $f : S \rightarrow \mathbb{R}^2$, el gradiente define una función (vectorial) $\nabla f : S \rightarrow \mathbb{R}^2$ dada por

$$(x, y) \longrightarrow \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right),$$

llamada habitualmente campo gradiente de la función f . En la siguiente sección analizaremos algunas propiedades de esta función.

1.1. El gradiente de una función

Sea $f : S \rightarrow \mathbb{R}$ una función de variable real, donde S es un abierto de $\mathbb{R}^n$. Supongamos que f tiene derivadas parciales $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq n$, continuas y consideramos el vector gradiente de f en $\mathbb{R}^n$ dado por

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

Es de señalar que si $\vec{u} = (u_1, \dots, u_n)$ es un vector unitario en $\mathbb{R}^n$, entonces la derivada direccional de f en la dirección $\vec{u}$ viene dada por la fórmula (6) y la del producto escala, es decir

$$D_{\vec{u}}f = \nabla f \cdot \vec{u} = \|\nabla f\| \cos(\angle(\nabla f, \vec{u})), \quad (7)$$

donde $\angle(\nabla f, \vec{u})$ denota el ángulo entre ∇f y $\vec{u}$. En consecuencia, tenemos las siguientes desigualdades que acotan los posibles valores de la derivada direccional de una función:

$$-\|\nabla f\| \leq D_{\vec{u}}f \leq \|\nabla f\|. \quad (8)$$

Dicho de otro modo, el vector gradiente apunta en la dirección y sentido de la máxima tasa de variación, y esta tasa máxima es $\|\nabla f\|$. Análogamente, la función f decrece más rápidamente en el sentido opuesto al gradiente de f , es decir, en la dirección $-\nabla f$. La tasa de variación de máximo decrecimiento de f es $-\|\nabla f\|$. Por otro lado, la derivada direccional es 0 en las direcciones ortogonales al gradiente. Para afianzar estas ideas mostramos cómo usarlas en unos ejemplos.

Ejemplo 1.2. Sea $f(x, y) = x^4 y^{-2}$ una función definida en $\mathbb{R}^2 \setminus \{y = 0\}$, y $P = (2, 1)$. Halla el vector unitario que apunte en la dirección y sentido de la máxima tasa de crecimiento de f en P .

Para calcular dicha tasa calcularemos el gradiente de la función y lo evaluaremos en P . Después tomaremos el vector unitario $\vec{u}$ correspondiente. Los cálculos son los siguientes.

$$\nabla f = (4x^3 y^{-2}, -2x^4 y^{-3}), \quad \nabla f(P) = (32, -32), \quad \vec{u} = (1/\sqrt{2}, 1/\sqrt{2}).$$

Ejemplo 1.3. La altitud de una montaña en variables (x, y) viene dada por la función

$$h(x, y) = 2500 + 100(x + y^2)e^{-0,3y^2},$$

donde x, y vienen dadas en unidades de 100 metros (m).

1. Halla la derivada direccional de f en el punto $P = (-1, -1)$ en la dirección del vector unitario $\vec{u}$ que forma un ángulo $\theta = \frac{\pi}{4}$ con el gradiente.
2. ¿Cuál es la interpretación de la derivada anterior?

La respuesta al primer apartado se halla usando la fórmula (7). En efecto,

$$\nabla f(-1, -1) = (74, -148), \quad D_{\vec{u}}f(-1, -1) = \|(74, -148)\| \cos\left(\frac{\pi}{4}\right) \simeq 165,5/\sqrt{2} \simeq 117.$$

La interpretación de esta derivada es sencilla. Si nos situamos en la montaña en el punto $(-1, -1, 2500)$, y empezamos a movernos sobre la montaña de forma que nuestro desplazamiento horizontal se realice en la dirección $\vec{u}$, entonces la altitud aumenta a razón de 117 metros por cada 100 metros de desplazamiento horizontal, o también 1,17 m por cada metro.

Otra propiedad importante del campo gradiente es que es perpendicular a las curvas de nivel de la función. Este hecho es una consecuencia de la regla de la cadena que estudiaremos a continuación, véase Proposición 1.1. Resumimos lo expuesto en esta sección en el siguiente teorema.

Teorema 1.2. (Interpretación geométrica del gradiente) Sea $f : S \rightarrow \mathbb{R}$ una función de variable real, donde S es un abierto de $\mathbb{R}^2$. Supongamos que f tiene derivadas parciales $\frac{\partial f}{\partial x_i}$, $1 \leq i \leq 2$, continuas y que el vector gradiente de f en el punto $P \in S$ no es el vector nulo. Sea $\vec{u}$ un vector unitario formando un ángulo θ con $\nabla f(P)$. Entonces

$$D_{\vec{u}}f(P) = \|\nabla f(P)\| \cos(\theta).$$

Además, se verifica

- $\nabla f(P)$ apunta en la dirección y sentido de la tasa de variación máxima de crecimiento de f en P .
- $-\nabla f(P)$ apunta en la dirección y sentido de la tasa de variación máxima de decrecimiento de f en P .
- $\nabla f(P)$ es ortogonal a la curva de nivel de f en P .

Nota 1.3. Es un ejercicio interesante la formulación y verificación del resultado análogo al anterior para funciones $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^n$, que se propone al lector.

1.2. Regla de la cadena en varias variables

Recordemos en primer lugar que si una función de una variable, $y = f(x)$, y la variable x varía con el tiempo t , es decir $x = x(t)$, entonces la fórmula de derivación llamada “Regla de la cadena” proporciona el cálculo de la variación de y como función de t . En concreto:

$$\frac{dy}{dt} = \frac{df}{dt} = \frac{df}{dx} \cdot \frac{dx}{dt}.$$

Por ejemplo, si $y = \cos(3x^2+1)$, y la variable x es la función dada por $x = x(t) = 2t^3$, entonces $\dot{y} = \frac{dy}{dt} = -\sin(12t^6+1) \cdot (6t^2)$. Comenzaremos estudiando la generalización de esta fórmula en el caso de varias variables. Es la llamada “Regla de la cadena para trayectoria”.

Proposición 1.1. (*Regla de la cadena para una trayectoria*) Sea $F : S \rightarrow \mathbb{R}$ una función de varias variables reales con $S \subset \mathbb{R}^n$, y un punto $P \in S$. Supongamos dada una curva en S con punto inicial P , es decir una función $\vec{x}(t)$ derivable y con derivada continua definida en un intervalo $[a, b] \subset \mathbb{R}$ tal que $\vec{x}(a) = P$,

$$[a, b] \rightarrow S, \quad \vec{x}(t) = (x_1(t), \dots, x_n(t)). \quad (9)$$

Entonces, la derivada respecto de la variable t de la función compuesta $F \circ \vec{x}(t)$ se calcula mediante la fórmula:

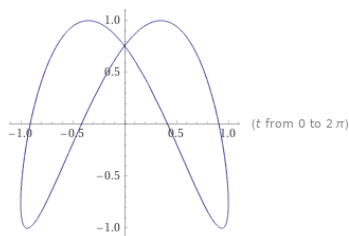
$$\frac{dF}{dt}(a) = \nabla F(P) \cdot \frac{d\vec{x}}{dt}(a) = \frac{\partial F}{\partial x_1}(P) \frac{dx_1}{dt}(a) + \dots + \frac{\partial F}{\partial x_n}(P) \frac{dx_n}{dt}(a). \quad (10)$$

La proposición anterior en el caso de dos variables (x, y) admite la siguiente formulación:

$$\frac{dF}{dt} = \nabla F \cdot \frac{d\vec{x}}{dt} = \frac{\partial F}{\partial x} \frac{dx}{dt} + \frac{\partial F}{\partial y} \frac{dy}{dt}. \quad (11)$$

Ilustraremos su aplicación sobre un ejemplo.

Ejemplo 1.4. Consideremos la función $F(x, y) = 20 + 10e^{-0.3(x^2+y^2)}$ que modeliza la función temperatura sobre una placa (en $^{\circ}\text{C}$), y la trayectoria $\vec{c} = (\cos(t-2), \sin(2t))$ (t en segundos), cuya gráfica se puede ver en la siguiente figura:



Entonces

$$\nabla F = \left(-6xe^{-0,3(x^2+y^2)}, -6ye^{-0,3(x^2+y^2)} \right), \quad \frac{d\vec{c}}{dt} = (-\sin(t-2), 2\cos(2t)) .$$

En consecuencia,

$$\frac{dF}{dt} = 6x \sin(t-2)e^{-0,3(x^2+y^2)} - 12y \cos(2t)e^{-0,3(x^2+y^2)} .$$

Si nos interesara la tasa de variación de la función temperatura F en el punto $P = \vec{c}(0) = (\cos(2), 0) \simeq (0,9, 0)$ al desplazarnos sobre la curva $\vec{c}(t)$ obtendríamos $\frac{dF}{dt}(0) \simeq 2,155^\circ\text{C/s}$.

Proposición 1.2. Sea $f : S \rightarrow \mathbb{R}$, con $S \subset \mathbb{R}^2$ abierto y f derivadas parciales continuas. Entonces, para cada punto $P \in S$, el vector $\nabla f(P)$ es perpendicular a la curva de nivel de f que pasa por P .

Demostración. En efecto, si $\vec{c}(t) = (x(t), y(t))$ es una parametrización de la curva de nivel de f que pasa por P , entonces, $f \circ \vec{c}(t) = \lambda$ para una cierta constante λ ; luego, por (11), se tiene que

$$\frac{df \circ \vec{c}}{dt} = \frac{d\lambda}{dt} = 0 \implies \nabla f(P) \cdot \vec{c}'(t) = 0,$$

lo que da el resultado, ya que $\vec{c}'(t)$ es el vector tangente a la curva de nivel λ . $\square$

La versión general de la Regla de la cadena (no sólo para una trayectoria) establece el siguiente resultado.

Teorema 1.3. Regla de la cadena para funciones. Sea $f : S \rightarrow \mathbb{R}$ una función definida en un conjunto abierto $S \subset \mathbb{R}^n$. Supongamos que cada función coordenada x_i es a su vez una función diferenciable de un conjunto de m variables reales, $\underline{t} = (t_1, \dots, t_m)$, para $i = 1, \dots, n$. Si denotamos por $\vec{x}(\underline{t}) = (x_1(\underline{t}), \dots, x_n(\underline{t}))$, entonces tenemos la siguiente fórmula de derivación:

$$\frac{\partial f \circ \vec{x}}{\partial t_k}(\underline{t}) = \nabla f \cdot \left(\frac{\partial x_1}{\partial t_k}, \dots, \frac{\partial x_n}{\partial t_k} \right) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \frac{\partial x_i}{\partial t_k} \quad (12)$$

Ilustraremos con un ejemplo cómo se aplica la fórmula (12).

Ejemplo 1.5. Consideremos la función $f(x, y) = x^2 + y^2 + x$, y el cambio de coordenadas a polares dado por

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Se pide calcular la variación de f con respecto a la variable radio r y también respecto a la variable angular θ . Para esto aplicaremos la fórmula (12) para las variables

$$(x_1, x_2) = (x, y), \quad (t_1, t_2) = (r, \theta).$$

En consecuencia

$$\frac{\partial f}{\partial r} = (2x + 1, 2y) \cdot (\cos \theta, \sin \theta) = (2r \cos \theta + 1, 2r \sin \theta) \cdot (\cos \theta, \sin \theta) = 2r + \cos \theta,$$

$$\frac{\partial f}{\partial \theta} = (2x + 1, 2y) \cdot (-r \sin \theta, r \cos \theta) = (2r \cos \theta + 1, 2r \sin \theta) \cdot (-r \sin \theta, r \cos \theta) = -r \sin \theta.$$

1.2.1. Derivación de funciones definidas de forma implícita

Supondamos dada una superficie Γ en $\mathbb{R}^3$ definida por la ecuación $F(x, y, z) = 0$. Supongamos que cerca de un punto P de Γ la coordenada z se puede poner como función de (x, y) , es decir $z = z(x, y)$, en este caso, $\frac{\partial F}{\partial x} \neq 0$, y se dan las siguientes fórmulas, llamadas de *Fórmulas de derivación en implícitas*

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \quad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}, \quad (13)$$

donde hemos usado las notaciones $F_x = \frac{\partial F}{\partial x}$, $F_y = \frac{\partial F}{\partial y}$, $F_z = \frac{\partial F}{\partial z}$. La verificación de la fórmula anterior es fácil: basta aplicar la regla de la cadena, (12), a la ecuación $F(x, y, z) = 0$. Mostraremos su uso en el siguiente ejemplo.

Ejemplo 1.6. Si consideramos el hiperboloide de una hoja de ecuación

$$F(x, y, z) = x^2 + y^2 - 2z^2 + 12x - 8z - 4 = 0,$$

entonces las variaciones $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$ en el punto $P = (1, 1, 1)$, vienen dadas por

$$\frac{\partial z}{\partial x}(P) = -\frac{2x + 12}{-4z - 8} \bigg|_P = \frac{7}{6}, \quad \frac{\partial z}{\partial y}(P) = -\frac{2y}{-4z - 8} \bigg|_P = \frac{1}{6}.$$

1.3. Diferenciabilidad y continuidad de funciones en varias variables

Consideremos una función $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Diremos que f es **diferenciable en un punto P_0 de $\mathbb{R}^n$** , si y sólo si existe una matriz T de tamaño $m \times n$ tal que

$$\lim_{X \rightarrow P_0} \frac{\|f(X) - [f(P_0) + T(\vec{P_0 X})]\|}{\|\vec{P_0 X}\|} = 0. \quad (14)$$

Llamaremos a la matriz T la derivada o diferencial de f en P_0 , y la denotaremos por $D_{P_0}f$, o también $Df(P_0)$. Se da el siguiente resultado.

Teorema 1.4. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en un punto $P_0 \in \mathbb{R}^n$. Entonces todas las derivadas parciales de f existen en P_0 y se tiene que la matriz $T = (t_{ij})$ de tamaño $m \times n$ que satisface la condición (14) es*

$$T = (t_{ij}) = \left(\frac{\partial f_i}{\partial x_j}(P_0) \right),$$

donde $f = (f_1, \dots, f_m)$ son las componentes de f y $(x_1, \dots, x_n)$ es un sistema de coordenadas en $\mathbb{R}^n$.

Demostración. Véase [3], pag. 158. □

En el caso de funciones de varias variables, también se da el hecho de que la diferenciabilidad de f en un punto implica la continuidad en ese punto. En concreto, usando el siguiente resultado se puede demostrar el teorema 1.5.

Lema 1.1. *Véase [3], pag 162.*

Teorema 1.5. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en un punto $P_0 \in \mathbb{R}^n$. Entonces f es continua en P_0 , y además existe una constante $M > 0$ tal que*

$$\|f(X) - f(P_0)\| < M\|X - P_0\|, \quad (15)$$

para X cerca de P_0 y $X \neq P_0$.

Demostración. Véase [3], pag. 158. □

Finalmente podemos concluir que, en el caso de que la función considerada f tenga derivadas parciales continuas, entonces f es diferenciable. En concreto, para el caso de dos variables tendríamos el siguiente enunciado.

Teorema 1.6. *Supongamos que $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable. Entonces todas las derivadas parciales existen, y la derivada direccional en un punto X en la dirección $\vec{u}$ está dada por la fórmula:*

$$D_{\vec{u}}f(X) = \nabla f(X) \cdot \vec{u} \quad (16)$$

Demostración. Consideremos la curva $\vec{c}(t) = X + t\vec{u}$. □

1.4. Aproximación de funciones y Teorema de Taylor

1.4.1. Plano tangente a una superficie en un punto. Aproximación lineal

Consideremos dada una superficie en $\mathbb{R}^3$ definida por una ecuación $F(x, y, z) = 0$, es decir $\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 \mid F(x, y, z) = 0\}$. Consideremos un punto $P_0 = (x_0, y_0, z_0)$ de dicha superficie, es decir $F(x_0, y_0, z_0) = 0$. Entonces, se define *el plano tangente a $\mathcal{S}$ en P_0* como el plano definido por la ecuación:

$$\frac{\partial F}{\partial x}(P_0)(x - x_0) + \frac{\partial F}{\partial y}(P_0)(y - y_0) + \frac{\partial F}{\partial z}(P_0)(z - z_0) = 0 . \quad (17)$$

Un caso particular importante es el de considerar las superficies que son gráficas de funciones de dos variables, $f(x, y)$. En este caso $F(x, y, z) = f(x, y) - z$, y la ecuación anterior tiene una fórmula más sencilla:

$$\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) - (z - z_0) = 0 , \quad (18)$$

donde $z_0 = f(x_0, y_0)$.

Ejemplo 1.7. 1. Si $F(x, y, z) = x^2 + y^2 + z^2 - 1$, entonces el plano tangente a la superficie $F = 0$ en el punto $P_0 = (1/\sqrt{2}, 1/\sqrt{2}, 0)$ es el plano

$$\frac{2}{\sqrt{2}}(x - \frac{1}{\sqrt{2}}) + \frac{2}{\sqrt{2}}(y - \frac{1}{\sqrt{2}}) = 0 . \quad (19)$$

2. Si $f(x, y) = x^2 - y^2$ y consideramos la superficie que da su gráfica, entonces, por ejemplo, el plano tangente a la gráfica en el punto (de la gráfica $P_0 = (1, 1, 0)$) es

$$2(x - 1) - 2(y - 1) - z = 0 . \quad (20)$$

Una utilidad de la ecuación del plano tangente a una gráfica de la forma $z = f(x, y)$ es para construir la aproximación lineal de la función f alrededor del punto (x_0, y_0) . Se tiene entonces la siguiente fórmula:

$$f(x, y) \simeq f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0), \quad (21)$$

fórmula que recibe el nombre de *aproximación lineal de f alrededor del punto (x_0, y_0)* .

Ejemplo 1.8. Si $f(x, y) = x^2 - y^2$, y consideramos el punto $P = (1, 1)$, la aproximación lineal de f alrededor del punto P es

$$f(x, y) \simeq 0 + 2(x - 1) - 2(y - 1) = 2x - 2y. \quad (22)$$

1.4.2. Derivadas de orden superior. Teorema de Taylor

Las derivadas parciales de orden superior son las derivadas parciales de las derivadas parciales de una función f . Así, si $f : S \rightarrow \mathbb{R}$ es una función de dos variables, por ejemplo (x, y) , las derivadas parciales de segundo orden son las derivadas parciales de $f_x = \frac{\partial f}{\partial x}$ y de $f_y = \frac{\partial f}{\partial y}$; es decir,

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

En las fórmulas anteriores se aprecia la importancia del orden de derivación iterada a la hora de calcular las *derivadas cruzadas* f_{yx} , f_{xy} . Sin embargo, para funciones muy generales tal ambigüedad desaparece debido al , también conocido como teorema de Schwarz.

Teorema 1.7. Si f_{yx} y f_{xy} son continuas en una bola B , entonces $f_{yx} = f_{xy}$ en cada punto de B .

Ilustraremos el resultado anterior sobre un ejemplo.

Ejemplo 1.9. Consideremos $f(x, y) = e^{x/y}$. Entonces $f_x = \frac{1}{y}e^{x/y}$, $f_y = -\frac{x}{y^2}e^{x/y}$. Entonces las derivadas segundas de f son:

$$f_{xx} = \frac{1}{y^2}e^{x/y}, \quad f_{yx} = \left(-\frac{1}{y^2} - \frac{x}{y^3}\right)e^{x/y},$$

$$f_{yy} = \left(\frac{2x}{y^3} + \frac{x}{y^4}\right)e^{x/y}, \quad f_{xy} = \left(-\frac{x}{y^3} - \frac{1}{y^2}\right)e^{x/y},$$

lo que comprueba la igualdad $f_{yx} = f_{xy}$.

Las derivadas de orden superior son un elemento importante de las funciones para construir sus aproximaciones polinómicas. En concreto, esto sucede con el teorema de Taylor 1.8 para funciones en varias variables.

La **notación** que usaremos es la siguiente, llamada notación de multiíndice. Si $\alpha = (\alpha_1, \dots, \alpha_n)$ es un vector de números naturales de tamaño n y $\mathbf{v} = (v_1, \dots, v_n)$, tomamos la siguiente notación:

$$\begin{aligned} |\alpha| &= \alpha_1 + \dots + \alpha_n, \\ \alpha! &= \alpha_1! \dots \alpha_n!, \\ \mathbf{v}^\alpha &= v_1^{\alpha_1} \dots v_n^{\alpha_n}. \end{aligned}$$

Además, si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es una función cuyas derivadas parciales de cualquier orden son diferenciables (para que no haya problema con el orden de las parciales), tomamos la notación:

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial \mathbf{x}^\alpha} = \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Con esta notación, el teorema de Taylor queda de la siguiente forma.

Teorema 1.8. (*Teorema de Taylor para funciones de varias variables*)

Sea $f : S \rightarrow \mathbb{R}$ una función. Supongamos que f y todas sus derivadas parciales de cualquier orden diferenciables en un punto $\mathbf{a} \in \mathbb{R}^n$ (por el orden de las parciales). Entonces, si $\mathbf{x} = (x_1, \dots, x_n)$, existen funciones $E_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$ para cada α con $|\alpha| = k$ tal que

$$\begin{aligned} f(\mathbf{x}) &= P_{k,\mathbf{a}}(\mathbf{x}) + \sum_{|\alpha|=k} E_\alpha(\mathbf{x})(\mathbf{x} - \mathbf{a})^\alpha \\ &= \sum_{|\alpha| \leq k} \frac{D^\alpha f(\mathbf{a})}{\alpha!} (\mathbf{x} - \mathbf{a})^\alpha + \sum_{|\alpha|=k} E_\alpha(\mathbf{x})(\mathbf{x} - \mathbf{a})^\alpha, \end{aligned}$$

donde además

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} E_\alpha(\mathbf{x}) = 0$$

para todos los E_α .

En el teorema anterior, $P_{k,\mathbf{a}}(\mathbf{x})$ se llama **el polinomio de Taylor de orden k de f en el punto $\mathbf{a}$** , y $E_\alpha(\mathbf{x})$ son las **funciones error** (que son muy pequeños cuando $\mathbf{x}$ está muy cerca del punto $\mathbf{a}$).

Dada la complejidad de la fórmula anterior la explicitaremos para el caso particular de funciones de dos variables $n = 2$ y orden de diferenciabilidad $m = 3$. En este caso el teorema anterior se puede reformular como sigue.

Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función. Supongamos que f y todas sus derivadas parciales son diferenciables en $\mathbf{a} = (x_0, y_0)$. Entonces, el polinomio de Taylor de orden 2 de f en $\mathbf{a}$ es

$$P_{2,\mathbf{a}}(x, y) = f(\mathbf{a}) + \frac{\partial f}{\partial x}(\mathbf{a})(x - x_0) + \frac{\partial f}{\partial y}(\mathbf{a})(y - y_0) \\ + \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) \frac{(x - x_0)^2}{2!} + \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a})(x - x_0)(y - y_0) + \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \frac{(y - y_0)^2}{2!}$$

y el polinomio de Taylor de orden 3 de f en $\mathbf{a}$ es, escribiendo $v_1 = x - x_0$ y $v_2 = y - y_0$,

$$P_{3,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \frac{\partial f}{\partial x}(\mathbf{a})v_1 + \frac{\partial f}{\partial y}(\mathbf{a})v_2 \\ + \frac{\partial^2 f}{\partial x^2}(\mathbf{a}) \frac{v_1^2}{2!} + \frac{\partial^2 f}{\partial x \partial y}(\mathbf{a})v_1v_2 + \frac{\partial^2 f}{\partial y^2}(\mathbf{a}) \frac{v_2^2}{2!} \\ + \frac{\partial^3 f}{\partial x^3}(\mathbf{a}) \frac{v_1^3}{3!} + \frac{\partial^3 f}{\partial x^2 \partial y}(\mathbf{a}) \frac{v_1^2v_2}{2!} + \frac{\partial^3 f}{\partial x \partial y^2}(\mathbf{a}) \frac{v_1v_2^2}{2!} + \frac{\partial^3 f}{\partial y^3}(\mathbf{a}) \frac{v_2^3}{3!}.$$

La fórmula de P_2 proporciona **la aproximación cuadrática de la función f cerca del punto (x_0, y_0)** , y se puede reescribir como

$$P_{2,(x_0,y_0)}(x, y) = f(x_0, y_0) + \nabla f(\mathbf{a}) \cdot (x - x_0, y - y_0) + \\ \frac{1}{2}(x - x_0, y - y_0) \begin{pmatrix} f_{xx}(\mathbf{a}) & f_{yx}(\mathbf{a}) \\ f_{xy}(\mathbf{a}) & f_{yy}(\mathbf{a}) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \quad (23)$$

La matriz 2×2 de la fórmula anterior se llama **la matriz Hessiana de f** y será fundamental para el estudio de los puntos extremos de la función f . Se denota por $Hess f$, y si la evaluamos en un punto $\mathbf{a}$, entonces escribiremos $Hess_{\mathbf{a}} f$.

Ejemplo 1.10. *Calcula la fórmula de aproximación cuadrática (23) para la función $f(x, y) = x^3 + y^3 - x$ en el punto $P = (1, -1)$.*

En primer lugar calculamos las derivadas necesarias:

$$\nabla f = (3x^2 - 1, 3y^2), \quad Hess f = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}.$$

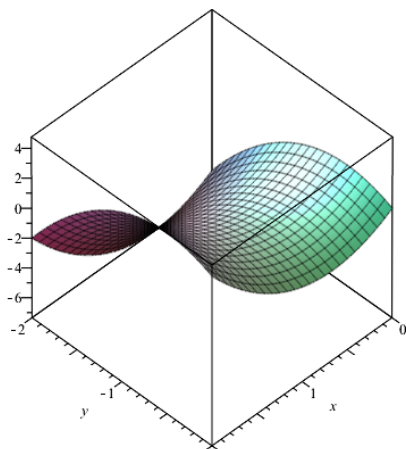
Después evaluamos en el punto dado:

$$f(P) = -1, \quad \nabla f(P) = (2, 3), \quad Hess_P f = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix},$$

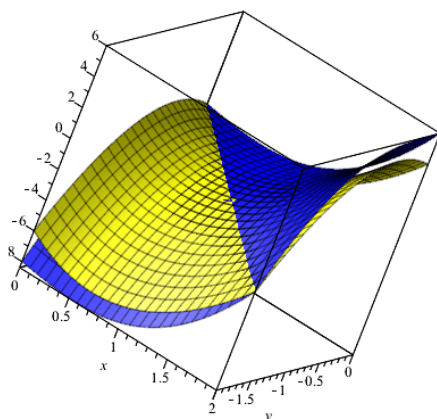
lo que da la aproximación cuadrática:

$$P_{2,P}(x, y) = -1 + 2(x-1) + 3(y+1) + \frac{1}{2}(x-1, y+1) \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} \begin{pmatrix} x-1 \\ y+1 \end{pmatrix} = 3x^2 - 3y^2 - 4x - 3y.$$

Observa que la superficie $z = 3x^2 - 3y^2 - 4x - 3y$ es un parabolide hiperbólico y tiene la representación gráfica siguiente



y proporciona una aproximación cuadrática local de la función f cerca del punto $P = (1, -1)$. Por ejemplo, usando el comando `plot3d([x^3+y^3-x, 3x^2-3y^2-4x-3y], x = 0 .. 2, y = -2 .. 0, color = [blue, yellow])` de Maple podemos representar ambas gráficas y se observa lo siguiente:



1.5. Máximos y mínimos de funciones de varias variables

La optimización es el proceso de hallar los valores extremos de una función. Esto consiste en hallar por ejemplo los puntos donde se alcanzan valores máximos y mínimos de la gráfica sobre un cierto dominio S . Análogamente al caso de una variable, es importante distinguir entre extremos locales y globales. Para tratar este tema necesitaremos establecer algunas definiciones básicas.

Definición 1.3. Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tiene un extremo local en un punto $P \in \mathbb{R}^n$ si existe una bola centrada en P , $B(P; \epsilon) \subset S$ tal que se da alguna de las dos situaciones siguientes

1. Para todo punto $\underline{x} \in B(P; \epsilon)$ se verifica la desigualdad $f(\underline{x}) \leq f(P)$. Situación llamada de **máximo local en P** .
2. Para todo punto $\underline{x} \in B(P; \epsilon)$ se verifica la desigualdad $f(\underline{x}) \geq f(P)$. Situación llamada de **mínimo local en P** .

Definición 1.4. Un punto $P \in \mathbb{R}^n$ se dice un punto crítico de f si el gradiente de f en P es el vector nulo, es decir $\nabla f(P) = \vec{0}$, o alguna derivada parcial de f no existe en el punto P .

Los conceptos anteriores están relacionados. El siguiente teorema (atribuido a Fermat) establece su conexión.

Teorema 1.9. Si f tiene un máximo o un mínimo local en un punto P , entonces P es un punto crítico de f .

Por simplicidad haremos la demostración en el caso de funciones de dos variables, aunque el resultado es cierto para cualquier número n de variables.

Demostración. Supongamos en primer lugar que f tiene un mínimo local en $P = (x_0, y_0)$. Entonces en una cierta bola B centrada en P se tiene que $f(x, y) \geq f(P)$. Entonces la función $g(x) := f(x, y_0)$ tiene un mínimo en $x = x_0$, y por tanto $g'(x_0) = 0$, puesto que $g'(x) = \frac{\partial f}{\partial x}(x, y_0)$. Luego $\frac{\partial f}{\partial x}(x_0, y_0) = 0$. Además, si definimo la función auxiliar $h(y) := f(x_0, y)$, se tiene que $h'(y) = \frac{\partial f}{\partial y}(x_0, y)$, y entonces h tiene un mínimo en y_0 . Por esto $h'(y_0) = 0$, lo que implica $\frac{\partial f}{\partial y}(x_0, y_0) = 0$. En consecuencia $\nabla f(P) = \vec{0}$.

Un razonamiento análogo se puede hacer para demostrar el enunciado en el caso de que P se un máximo local de f . Se propone al lector completar los detalles de la demostración en este caso. $\square$

Ilustraremos el método de cálculo de extremos locales mediante un ejemplo.

Ejemplo 1.11. Sea $f(x, y) = 11x^2 - 2xy + 2y^2 + 3y$. Calcula los puntos críticos de f .

En primer lugar calcularemos el gradiente de la función:

$$\nabla f = (22x - 2y, -2x + 4y + 3),$$

para imponer la condición $\nabla f = \vec{0}$. Lo que nos proporciona el sistema de ecuaciones

$$22x - 2y = 0, -2x + 4y + 3 = 0,$$

cuya única solución es $x = -1/14$, $y = -11/14$. Luego el extremo es el punto $P = (-1/14, -11/14)$. El valor de la función en este punto es $f(P) = -33/28$.

En el caso de funciones de dos variables se puede establecer un criterio análogo al que se tiene en el caso de las funciones de una variable para determinar el tipo de extremo de la función. Lo enunciaremos a continuación.

Teorema 1.10. (*Test de las segundas derivadas*)

Sea P un punto crítico de una función $f(x, y)$. Supongamos que existen las derivadas segundas f_{xx} , f_{yx} , f_{yy} y son continuas cerca de P . Consideremos la matriz Hessiana de f en P :

$$Hess_P f = \begin{pmatrix} f_{xx}(P) & f_{yx}(P) \\ f_{xy}(P) & f_{yy}(P) \end{pmatrix}.$$

Entonces:

1. Si $\det(Hess_P f) > 0$, entonces
 - Si $f_{xx}(P) > 0$, el punto P es un mínimo local de f .
 - Si $f_{xx}(P) < 0$, el punto P es un máximo local de f .
2. Si $\det(Hess_P f) < 0$, entonces P es un punto de silla.
3. Si $\det(Hess_P f) = 0$, el criterio no decide.

Es de señalar que si aplicamos este criterio a la función del ejemplo 1.11 obtenemos que $P = (-1/14, -11/14)$ es un mínimo de la función ya que $D = 84 > 0$ y también $f_{xx}(P) = 22 > 0$. Quedaría por determinar si este extremo es global o solamente local.

Teorema 1.11. Sea f una función continua sobre un dominio cerrado y acotado $\mathcal{D} \subset \mathbb{R}^2$. Entonces

- La función f alcanza un valor máximo y uno mínimo en $\mathcal{D}$.
- Los extremos ocurren o bien en los puntos críticos del interior de $\mathcal{D}$, o bien en los puntos de la frontera de $\mathcal{D}$.

Ejemplo 1.12. Halla el valor máximo de la función $f(x, y) = 2x + y - 3xy$ sobre el cuadrado unidad $\mathcal{D} = [0, 1] \times [0, 1]$.

1. Halla la función a maximizar.

En esta ocasión viene dada: $f(x, y) = 2x + y - 3xy$.

2. Examina los puntos críticos.

Se calcula el gradiente de f , que es $\nabla f = (2-3y, 1-3x)$, y se iguala a cero. Nos queda un sistema de ecuaciones cuya solución única es $(x, y) = (1/3, 2/3)$. Si quisiéramos distinguir si es un máximo local o un mínimo local, calcularíamos el determinante del hessiano:

$$Hess_{(x,y)}f = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix},$$

que evaluado en el $(1/3, 2/3)$ tiene determinante negativo, por lo tanto es un punto de silla. Este paso no sería necesario en nuestro caso, puesto que buscamos el máximo global, y para ello solo necesitamos comparar los valores de los puntos críticos, $f(1/3, 2/3) = 2/3$.

3. Analiza los puntos en la frontera. Los puntos de la frontera son los lados del cuadrado, es decir:

- $\ell_1 = \{(t, 0) : t \in [0, 1]\}$,
- $\ell_2 = \{(0, t) : t \in [0, 1]\}$,
- $\ell_3 = \{(t, 1) : t \in [0, 1]\}$ y
- $\ell_4 = \{(1, t) : t \in [0, 1]\}$.

Ahora analizamos los valores de f en esos lados:

- En los puntos de ℓ_1 , f toma valores $f(t, 0) = 2t$, y como t varía entre 0 y 1, el máximo de este lado se alcanza en $(1, 0)$ y tiene valor $f(1, 0) = 2$.

- En los puntos de ℓ_2 , f toma valores $f(0, t) = t$, y como t varía entre 0 y 1, el máximo de este lado se alcanza en $(0, 1)$ y tiene valor $f(0, 1) = 1$.
- En los puntos de ℓ_3 , f toma valores $f(t, 1) = 2t + 1 - 3t = 1 - t$, y como t varía entre 0 y 1, el máximo de este lado se alcanza en $(0, 1)$ y tiene valor $f(0, 1) = 1$.
- En los puntos de ℓ_4 , f toma valores $f(1, t) = 2 + t - 3t = 2 - 2t$, y como t varía entre 0 y 1, el máximo de este lado se alcanza en $(1, 0)$ y tiene valor $f(1, 0) = 2$.

4. **Compara.** Entre los valores críticos de f y los máximos en los lados, el mayor de todos los valores se alcanza en $(1, 0)$, por lo tanto es el máximo global y su valor es $f(1, 0) = 2$.

1.5.1. Extremos condicionados de funciones

Teorema 1.12. Supongamos dadas dos funciones diferenciables $f(x, y)$, $g(x, y)$. Si f tiene un mínimo o un máximo local sobre la curva dada implícitamente por la ecuación $g(x, y) = 0$ en el punto $P = (x_0, y_0)$ tal que $\nabla g(P) \neq \vec{0}$, entonces existe un escalar λ tal que

$$\nabla f(P) = \lambda \nabla g(P) . \quad (24)$$

Ejemplo 1.13. Halla los valores críticos de la función $f(x, y) = 2x + 5y$ sobre la elipse de ecuación:

$$\left(\frac{x}{4}\right)^2 + \left(\frac{y}{3}\right)^2 = 1.$$

Procederemos de la manera siguiente:

1. **Escribe la ecuación de Lagrange** (24). Primero, tenemos que identificar las funciones que intervienen. La f está dada, pero la g es $g(x, y) = \left(\frac{x}{4}\right)^2 + \left(\frac{y}{3}\right)^2 - 1$, puesto que queremos una ecuación del tipo $g(x, y) = 0$. El gradiente de f es $\nabla f(x, y) = (2, 5)$, y el de g es $\nabla g(x, y) = (x/8, 2y/9)$. La ecuación de Lagrange es:

$$(2, 5) = \lambda(x/8, 2y/9)$$

2. **Aisla λ como función de (x, y) .**

Esto es fácil, puesto que de la ecuación de Lagrange tenemos tanto

$$2 = \frac{\lambda x}{8}$$

como

$$5 = \frac{2\lambda y}{9},$$

y despejamos λ , que sería

$$\lambda = \frac{16}{x}$$

y también

$$\lambda = \frac{45}{2y}.$$

3. **Halla x_0 e y_0 usando la restricción.** De las dos ecuaciones anteriores obtenemos que

$$\frac{16}{x} = \frac{45}{2y},$$

o, despejando x ,

$$x = \frac{32y}{45}.$$

Ahora es cuando usamos nuestra restricción $g(x, y) = 0$, que como sabemos que $x = \frac{32y}{45}$ queda

$$\begin{aligned} g\left(\frac{32y}{45}, y\right) &= 0, \\ \left(\frac{\frac{32y}{45}}{4}\right)^2 + \left(\frac{y}{3}\right)^2 - 1 &= 0, \\ \frac{64y^2}{45^2} + \frac{y^2}{9} - 1 &= 0. \end{aligned}$$

Si despejamos la y , es $y = \pm \frac{45}{17}$, que tiene respectivas $x = \pm \frac{32}{17}$. Además f toma los valores en esos puntos

$$f\left(\frac{45}{17}, \frac{8}{17}\right) = \frac{130}{17}$$

y

$$f\left(-\frac{45}{17}, -\frac{8}{17}\right) = -\frac{130}{17}.$$

Por lo tanto, los valores críticos son $\left(\frac{15}{17}, \frac{8}{17}\right)$ y $\left(-\frac{15}{17}, -\frac{8}{17}\right)$ que son, respectivamente, un máximo y un mínimo globales.

Referencias

- [1] T. M. Apostol. *Análisis Matemático*. Reverté, 1982.
- [2] E. Hernández, M. J. Vázquez, and M. A. Zurro. *Álgebra Lineal y Geometría*. Pearson, 2012.
- [3] J.E. Marsden and A.J. Tromba. *Cálculo Vectorial*. Addison-Wesley Iberoamericana, 2004.
- [4] J. Rogawski. *Cálculo. Varias Variables*. Reverté, 2012.

Cálculo integral en $\mathbb{R}^2$ y $\mathbb{R}^3$

Prof. Roberto Giménez Conejero

20 de febrero de 2025

1. FUNCIONES DE DOS VARIABLES

Considera una función de dos variables $f : \Omega \rightarrow \mathbf{R}$, donde Ω es un conjunto contenido en el plano $\mathbf{R}^2$.

- Recintos de tipo I. Si Ω es un recinto que puede ser descrito de la forma

$$\Omega = \{(x, y) \in \mathbf{R}^2 : a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\} .$$

Entonces

$$\int_{\Omega} f(x, y) dx dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx .$$

- Recintos de tipo II. Si Ω es un recinto que puede ser descrito de la forma

$$\Omega = \{(x, y) \in \mathbf{R}^2 : c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\} .$$

Entonces

$$\int_{\Omega} f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy .$$

Coordenadas polares en $\mathbf{R}^2$. Cada punto $P = (x, y)$ de $\mathbf{R}^2$ tiene asignado de forma única una pareja de números reales¹ (r, θ) de forma que

$$x = r \cos \theta, \quad y = r \sin \theta .$$

Esta pareja (r, θ) recibe el nombre de coordenadas polares del punto P .

¹Con $r \geq 0$ y $0 \leq \theta < 2\pi$.

Teorema del cambio de variable a coordenadas polares. Si Ω es un recinto de $\mathbf{R}^2$ descrito de la forma

$$\Omega = \{(r \cos \theta, r \sin \theta) : \alpha \leq \theta \leq \beta, g_1(\theta) \leq r \leq g_2(\theta)\},$$

entonces

$$\int_{\Omega} f(x, y) dx dy = \int_{\alpha}^{\beta} \left(\int_{g_1(\theta)}^{g_2(\theta)} f(r \cos \theta, r \sin \theta) r dr \right) d\theta.$$

2. FUNCIONES DE TRES VARIABLES

Considera una función de tres variables $f : D \rightarrow \mathbf{R}$, donde D es un conjunto contenido en el espacio $\mathbf{R}^3$.

- Recintos típicos. Si D es un recinto que puede ser descrito de la forma

$$D = \{(x, y, z) \in \mathbf{R}^3 : a \leq x \leq b, g_1(x) \leq y \leq g_2(x), h_1(x, y) \leq z \leq h_2(x, y)\}.$$

Entonces

$$\int_D f(x, y, z) dx dy dz = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} \left(\int_{h_1(x, y)}^{h_2(x, y)} f(x, y, z) dz \right) dy \right) dx.$$

Coordenadas cilíndricas en $\mathbf{R}^3$. Cada punto $P = (x, y, z)$ de $\mathbf{R}^3$ tiene asignado de forma única una terna de números reales² (r, θ, z) de forma que

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Esta pareja (r, θ, z) recibe el nombre de coordenadas cilíndricas del punto P .

Teorema del cambio de variable a coordenadas cilíndricas. Si D es un recinto de $\mathbf{R}^3$ descrito de la forma

$$\Omega = \{(r \cos \theta, r \sin \theta, z) : \alpha \leq \theta \leq \beta, g_1(\theta) \leq r \leq g_2(\theta), h_1(r, \theta) \leq z \leq h_2(r, \theta)\},$$

entonces

$$\int_D f(x, y, z) dx dy dz = \int_{\alpha}^{\beta} \left(\int_{g_1(\theta)}^{g_2(\theta)} \left(\int_{h_1(r, \theta)}^{h_2(r, \theta)} f(r \cos \theta, r \sin \theta, z) dz \right) r dr \right) d\theta.$$

²Con $r \geq 0$ y $0 \leq \theta < 2\pi$.

Coordenadas esféricas en $\mathbf{R}^3$. Cada punto $P = (x, y, z)$ de $\mathbf{R}^3$ tiene asignado de forma única una terna de números reales³ (r, θ, ϕ) de forma que

$$x = r \cos \theta \sin \phi \quad , \quad y = r \sin \theta \sin \phi \quad , \quad z = r \cos \phi \quad .$$

Esta pareja (r, θ, ϕ) recibe el nombre de coordenadas esféricas del punto P .

Teorema del cambio de variable a coordenadas esféricas. Si D es un recinto de $\mathbf{R}^3$ descrito de la forma

$$D = \{(r \cos \theta \sin \phi, r \sin \theta \sin \phi, r \cos \phi) :$$

$$a_1 \leq \theta \leq a_2 \text{ , } g_1(\theta) \leq \phi \leq g_2(\theta) \text{ , } h_1(\theta, \phi) \leq r \leq h_2(\theta, \phi)\} \text{ ,}$$

entonces

$$\int_D f(x, y, z) dx dy dz = \int_{a_1}^{a_2} \left(\int_{g_1(\theta)}^{g_2(\theta)} \left(\int_{h_1(\theta, \phi)}^{h_2(\theta, \phi)} f(r \cos \theta \sin \phi, r \sin \theta \sin \phi, r \cos \phi) r^2 \sin \phi \, dr \right) d\phi \right) d\theta \text{ .}$$

³Con $r \geq 0$, $0 \leq \theta < 2\pi$ y $0 \leq \phi < \pi$.

1. Integración en varias variables

El cálculo de una integral múltiple (en varias variables) se reduce a ir calculando integrales de una variable en el orden especificado. Intuitivamente, al igual que en el caso de una variable, las integrales múltiples calculan volúmenes. Veamos un ejemplo en el caso de dos variables:

Ejemplo 1. Supongamos que queremos calcular la integral múltiple

$$I = \int_0^1 \int_2^3 (6x + 6y^2) dx dy$$

donde las variables x, y tienen de límites $[2, 3]$ y $[0, 1]$ respectivamente. Para realizar la integración, primero empezamos integrando la variable x , manteniendo y constante,

$$I = \int_0^1 (3x^2 + 6y^2 x) \Big|_{x=2}^{x=3} dy,$$

y luego integramos sobre y ,

$$I = \int_0^1 (15 + 6y^2) dy = (15y + 2x^3) \Big|_{y=0}^{y=1} = 17.$$

En la integral anterior se puede comprobar que el orden de integración no importa, es decir,

$$\int_2^3 \int_0^1 (6x + 6y^2) dx dy = \int_0^1 \int_2^3 (6x + 6y^2) dy dx.$$

Sin embargo, observa que el límite de integración hay que ordenarlo también. Esto sucede gracias al teorema de *Fubini*, que no vamos a enunciar porque tiene unas hipótesis técnicas. En la práctica, cualquier función razonable cumple el teorema de Fubini.

Como decíamos anteriormente, la integral $\int_c^d \int_a^b f(x, y) dx dy$ es el volumen que la gráfica de $f(x, y)$ deja sobre el rectángulo $[a, b] \times [c, d]$ (si $|f| > 0$ en ese rectángulo). En general no haremos integrales sobre rectángulos, sino sobre regiones R del plano (o de $\mathbb{R}^3$). En ese caso, se puede escribir

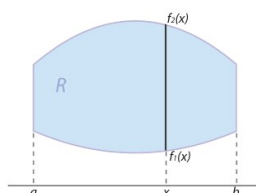
$$\iint_R f(x, y) dx dy.$$

De esta forma, cuando $f(x, y) = 1$, obtenemos:

$$\iint_R 1 dx dy = \text{Volumen del sólido que tiene base } R \text{ y altura } 1 = \text{Área de } R.$$

A veces, las regiones están descritas con funciones en las variables x o y :

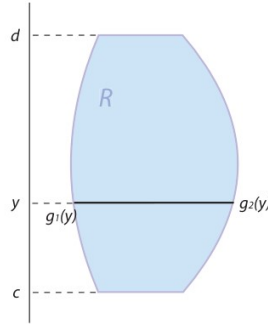
- Sugongamos una región R definida por las funciones $f_1(x)$ y $f_2(x)$, como por ejemplo



Debemos entonces integrar primero en y para tener en cuenta todas las secciones verticales, es decir

$$\iint_R F = \int_a^b \int_{f_1(x)}^{f_2(x)} F(x, y) dy dx$$

- Sugongamos una región R definida por las funciones $g_1(y)$ y $g_2(y)$, como por ejemplo



Debemos entonces integrar primero en x para tener en cuenta todas las secciones horizontales, es decir

$$\iint_R F = \int_c^d \int_{g_1(y)}^{g_2(y)} F(x, y) dx dy$$

A veces puede ocurrir también que las dos estrategias anteriores no funcionen, y tengamos que separar la región en trozos donde sí funcionen las estrategias anteriores. Esto es lo mismo que ocurre con las funciones definidas a trozos en una variable: hay que dividir la integral según los trozos y luego sumar todos los resultados.

Por último, como es evidente, se puede integrar en cuantas variables se quiera. Por ejemplo, se puede integrar en 3 variables, y el proceso es exactamente igual al anterior. En particular, se tiene que para la función $f(x, y, z) = 1$ y el sólido V ,

$$\int \int \int_V 1 dx dy dz = \text{Volumen del sólido } V.$$

De esta forma, si se sabe que un objeto tridimensional V tiene función de densidad $\rho(x, y, z)$ que depende del punto (x, y, z) donde se mida, se tiene que

$$\int \int \int_V \rho dx dy dz = \text{Masa del sólido } V,$$

que intuitivamente se entiende como masa=densidad·volumen.

En cualquier número de variables, si la función a integrar $f(x, y)$ es un producto de funciones $g(x)h(y)$, se puede ver fácilmente que

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

Nota: Es común que, cuando el contexto está claro, se eviten varios símbolos de la integral. Por ejemplo, si está claro por el contexto que estamos en el plano, todas las siguientes notaciones pueden aparecer en un libro: $\int \int_R f(x, y) dx dy$, $\int_R f(x, y) dx dy$, $\int_R f$.

Vamos a motivar la siguiente sección dando un ejemplo, que es muy difícil de calcular sin un *cambio de variable*.

Ejemplo 2. Calculemos el área del círculo de radio ℓ y centrado en el origen:

$$R = \{x^2 + y^2 \leq \ell^2\}.$$

En primer lugar, para calcular un área de una región sencillamente tenemos que calcular

$$A = \int \int_R 1 dx dy.$$

Reexpresamos R ,

$$R = \{-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\},$$

de tal forma que queda como una de las regiones que escribíamos anteriormente:

$$A = \int_{-\ell}^{\ell} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy dx,$$

donde también sabemos que x varía en $[-\ell, \ell]$. Hacemos la primera integral iterada:

$$A = \int_{-\ell}^{\ell} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy dx = \int_{-\ell}^{\ell} 2\sqrt{1-x^2} dx.$$

Esta integral es complicada. Hay que saber que este tipo de integrales se resuelven con el cambio de variable $x = \ell \sin(u)$, $dx = \ell \cos(u) du$,

$$A = \int_{-\pi/2}^{\pi/2} 2\ell^2 \cos^2(u) du.$$

Además hay que usar el truco de tomar $2\cos^2(u) = \cos(2u) + 1$:

$$A = \int_{-\pi/2}^{\pi/2} (\ell^2 \cos(2u) + \ell^2) du = \ell^2 \left(\frac{\sin(2u)}{2} + u \right) \Big|_{u=-\pi/2}^{u=\pi/2} = \ell^2(0 + \pi) = \ell^2 \pi.$$

La integral, que ya ha necesitado un cambio de variable al llegar a la integración en solo una variable, sería muchísimo más simple cambiando de variables en $\mathbb{R}^2$ a coordenadas polares.

2. Cambio de variable

Por incontables razones (al menos dos: simplificar la integral y simplificar la región de integración), es conveniente a veces hacer un cambio de variable, como ejemplificábamos antes. En analogía con el caso de una variable, se tiene el siguiente resultado, que lo enunciamos para 2 variables pero vale para cualquier número de variables.

Proposición 1. Dada una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ integrable y una región R en $\mathbb{R}^2$, si $\phi(u, v) = (x(u, v), y(u, v))$ es un cambio de variable $(u, v) \mapsto (x, y)$ (es decir, una aplicación biectiva), entonces

$$\int \int_R f(u, v) du dv = \int \int_R f(x, y) |det D\phi|,$$

donde $D\phi$ es la matriz diferencial y a la cual se tiene que hacer el valor absoluto del determinante.

Veamos el ejemplo anterior:

Ejemplo 3. Calculemos el área del círculo de radio ℓ y centrado en el origen:

$$R = \{x^2 + y^2 \leq \ell^2\}.$$

Podemos cambiar a coordenadas polares, que suelen funcionar mejor cuando tenemos cosas redondas o aparecen senos y cosenos en la integral. Como

$$\begin{aligned} x &= r \cos(\theta) \\ y &= r \sin(\theta), \end{aligned}$$

esto nos define una función de cambio de variable $\phi(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$. Como anteriormente, calculamos

$$A = \int \int_R 1 dx dy.$$

Como queremos cambiar de variables, tenemos que calcular la diferencial de ϕ y su determinante:

$$\det D\phi = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Tendríamos que tomar también el valor absoluto, pero r ya toma valores positivos, así que es igual a su valor absoluto. Al hacer el cambio de variable, θ varía en $[0, 2\pi]$ (para cubrir todos los ángulos del círculo) y r varía en $[0, \ell]$. Queda

$$A = \int_0^{2\pi} \int_0^{\ell} r dr d\theta = \int_0^{2\pi} \frac{r^2}{2} \Big|_{r=0}^{r=\ell} d\theta = \int_0^{2\pi} \frac{\ell^2}{2} d\theta = \frac{\ell^2}{2} \theta \Big|_{\theta=0}^{\theta=2\pi} = \pi \ell^2.$$

Este cálculo ha sido mucho más sencillo.

Ahora damos los cambios típicos de coordenadas.

2.1. A polares

Como el cambio de variable está definido por $\phi(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$, su diferencial es

$$D\phi = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

A su vez, $|\det D\phi| = r$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y) dx dy = \int_R f(r \cos \theta, r \sin \theta) r dr d\theta.$$

2.2. A cilíndricas

El cambio a cilíndricas es como el de polares con, además, una z : $\phi(r, \theta, z) = (r \cos \theta, r \sin \theta, z) = (x, y, z)$

$$D\phi = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A su vez, $|\det D\phi| = r$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y, z) dx dy dz = \int_R f(r \cos \theta, r \sin \theta, z) r dr d\theta dz.$$

2.3. A esféricas

Como el cambio de variable está definido por $\phi(r, \phi, \theta) = (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) = (x, y, z)$, su diferencial es

$$D\phi = \begin{pmatrix} \cos \phi \sin \theta & -r \sin \phi \sin \theta & r \cos \phi \cos \theta \\ \sin \phi \sin \theta & r \cos \phi \sin \theta & r \sin \phi \cos \theta \\ \cos \theta & 0 & -r \sin \theta \end{pmatrix}.$$

A su vez, $|\det D\phi| = r^2 \sin \theta$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y, z) dx dy dz = \int_R f(r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) r^2 \sin \theta dr d\phi d\theta.$$

Ejemplo 4. Esto nos permite calcular el volumen de una esfera de radio ℓ ,

$$V = \int_R 1 dx dy dz,$$

donde R se puede describir como $\{x^2 + y^2 + z^2 \leq \ell^2\}$. Es mucho más sencillo cambiar a coordenadas esféricas, con las variables los valores $r \in [0, \ell]$, $\phi \in [0, 2\pi]$ y $\theta \in [0, \pi]$:

$$V = \int_0^\pi \int_0^{2\pi} \int_0^\ell r^2 \sin \theta dr d\phi d\theta = \int_0^{2\pi} d\phi \int_0^\ell r^2 dr \int_0^\pi \sin \theta d\theta,$$

porque la función que nos queda es un producto de funciones en cada variable. Cuando se calcula esto último da

$$V = \frac{4\pi}{3} \ell^3.$$

Nota: Existen unas [coordenadas esféricas para las esferas de cualquier dimensión](#), muy parecidas a las que hemos visto nosotros para 3 dimensiones. Eso permite calcular el [volumen de las bolas de \$n\$ dimensiones](#).

3. Repaso geométrico

En esta primera sección recordamos los conceptos geométricos básicos. A partir de ellos podremos definir nociones como la longitud y el trabajo de una fuerza.

3.1. Operaciones con vectores

Las dos operaciones relevantes para el cálculo multivariable son:

Producto escalar

Dados dos vectores $v, w \in \mathbb{R}^2$ de coordenadas $v = (v_1, v_2)$, $w = (w_1, w_2)$. La proyección de w sobre v es el vector de dirección v y módulo

$$w \cdot v := v_1 w_1 + v_2 w_2 \in \mathbb{R}.$$

En el caso de tener vectores en el espacio $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3) \in \mathbb{R}^3$, la proyección de w sobre v es el vector que tiene dirección v y módulo

$$w \cdot v := v_1 w_1 + v_2 w_2 + v_3 w_3 \in \mathbb{R}.$$

En ambos casos, el **número real** $w \cdot v$ se conoce como el producto escalar de v y w .

Dos vectores v, w son ortogonales si $w \cdot v = 0$.

Producto vectorial

Esta operación es exclusiva de vectores en el espacio $\mathbb{R}^3$. Dados dos vector $v, w \in \mathbb{R}^3$, el producto vectorial de v y w es el **vector** ortogonal a v y w de módulo el área con signo del paralelogramo generado por v y w .

En coordenadas, $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3)$, el producto vectorial de v y w es

$$\begin{aligned} v \times w &:= \det \begin{pmatrix} v_1 & w_1 & \vec{i} \\ v_2 & w_2 & \vec{j} \\ v_3 & w_3 & \vec{k} \end{pmatrix} = (v_2 w_3 - v_3 w_2) \cdot \vec{i} + (v_3 w_1 - v_1 w_3) \cdot \vec{j} + (v_1 w_2 - v_2 w_1) \cdot \vec{k} = \\ &= (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1). \end{aligned}$$

Notemos que $v \cdot (v \times w) = 0$ y $w \cdot (v \times w) = 0$.

3.2. Curvas

En una primera aproximación, una curva es un objeto unidimensional. Es decir, un conjunto de puntos de cierto espacio que cuando se recorre sólo se tiene la libertad de moverse en una dirección, con dos posibles sentidos. Un vehículo desplazándose a lo largo de una calle de un solo carril está restringido a ir hacia delante o hacia atrás: la calle puede curvarse pero sigue habiendo sólo una dirección. Un funambulista al que no se le permita saltar está obligado a moverse a lo largo de un objeto unidimensional, el cable donde se apoya. La calle por la que se desplaza el vehículo se dibuja mediante una curva en un mapa de la ciudad correspondiente, es lo que denominaremos una curva plana. El cable del funambulista vive dentro del 3-espacio y dibuja una curva en el espacio.

3.2.1. Curvas en el plano

Sea $I \subset \mathbb{R}$ un intervalo en la recta $\mathbb{R}$ de variable independiente t . Una curva en el plano $\mathbb{R}^2$ de variables (x, y) es la imagen de una aplicación diferenciable

$$r : I \subset \mathbb{R} \longrightarrow \mathbb{R}^2, \quad t \longmapsto r(t) = (x(t), y(t)).$$

Se dice que la función r parametriza la curva.

Ejemplos: 1. **Segmentos.** Sea $I = [0, 1]$ y $r(t) = (t, 0)$. La curva imagen es el segmento en el plano $\mathbb{R}^2$ que une los puntos $(0, 0)$ y $(1, 0)$. Si $I = [-1, 0]$ y $r(t) = (t, t + 1)$. La curva es el segmento que une los puntos $(-1, 0)$ y $(0, 1)$. En general es segmento que une los punto $p, q \in \mathbb{R}^2$ se puede parametrizar con $r(t) = t \cdot q + (1 - t) \cdot p$, $0 \leq t \leq 1$.

2. **Circunferencias.** Sea $I = [0, 2\pi]$ y $r(t) = (\cos t, \sin t)$. La curva es la circunferencia de radio 1 centrada en el origen $(0, 0)$. Para describir la circunferencia de radio a centrada en el origen podemos usar $I = [0, 2\pi]$ y $r(t) = (a \cos t, a \sin t)$.

3. Sea $I = [0, \pi]$ y $r(t) = (\cos t, \sin t)$. Entonces la curva es la parte de la circunferencia unidad centrada en el origen contenida en el semiplano superior $\{(x, y) : y \geq 0\}$.

4. **Elipses.** Sea C la elipse de semiejes $a, b \geq 0$:

$$C = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}.$$

Una posible parametrización es $r(t) = (a \cos t, b \sin t)$, $0 \leq t \leq 2\pi$.

La función $r(t)$ suele interpretarse como la posición de una partícula moviéndose a lo largo del tiempo t . Esto motiva la siguiente

Definición 1. La velocidad de una curva $r : I \rightarrow \mathbb{R}^2$ es el vector $r'(t) = (x'(t), y'(t))$.

Ejemplos: 1. **Segmentos.** Sea $r(t) = (t, 0)$. La velocidad es $r'(t) = (1, 0)$. Si la posición está descrita por la curva $r(t) = (t^2 + t, te^t)$, la velocidad es $r'(t) = (2t + 1, e^t + te^t)$. Para un segmento general $r(t) = t \cdot q + (1 - t) \cdot p$ la velocidad es constante igual a $r'(t) = q - p$.

2. **Circunferencias.** En el caso de la circunferencia unidad $r(t) = (\cos t, \sin t)$ la velocidad es $r'(t) = (-\sin t, \cos t)$. Si se tiene la circunferencia de radio a descrita por $r(t) = (a \cos t, a \sin t)$, la velocidad es $r'(t) = (-a \sin t, a \cos t)$.

3. **Elipses.** Sea C la elipse parametrizada por

$$r(t) = (a \cos t, b \sin t), \quad 0 \leq t \leq 2\pi.$$

La velocidad es $r'(t) = (-a \sin t, b \cos t)$.

3.2.2. Curvas en el espacio

La noción de curva sigue siendo la imagen de un intervalo $I \subset \mathbb{R}$ por una aplicación continua r . Para curvas planas la aplicación es $r : I \rightarrow \mathbb{R}^2$. Para una curva en el espacio sólo tenemos que admitir que el conjunto de llegada sea $\mathbb{R}^3$.

Una curva en el espacio $\mathbb{R}^3$ de variables (x, y, z) es la imagen de una aplicación diferenciable

$$r : I \subset \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \mapsto r(t) = (x(t), y(t), z(t)).$$

Decimos que la función r parametriza la curva. La definición es análoga al caso de curvas plana añadiendo una tercera componente $z(t)$.

La noción de velocidad es análogo que en el caso de curvas planas:

Definición 2. La velocidad de una curva $r : I \rightarrow \mathbb{R}^3$ es el vector $r'(t) = (x'(t), y'(t), z'(t))$.

Siendo más precisos, la velocidad de una curva parametrizada por $r(t) = (x(t), y(t), z(t))$ en el punto $r(t_0)$ es el vector $r'(t_0) = (x'(t_0), y'(t_0), z'(t_0))$.

Ejemplos: 1. **Segmentos.** Consideremos los puntos $(1, 0, 2), (-3, 2, 4) \in \mathbb{R}^3$. El segmento que los une es una curva C parametrizada por

$$r : [0, 1] \rightarrow \mathbb{R}^3, \quad r(t) = t \cdot (-3, 2, 4) + (1 - t) \cdot (1, 0, 2) = (-4t + 1, 2t, 2t + 2), \quad 0 \leq t \leq 1.$$

Observemos que efectivamente $r(0) = (1, 0, 2)$ y $r(1) = (-3, 2, 4)$. En general dados dos puntos $p, q \in \mathbb{R}^3$ en el espacio, el segmento que los une se parametriza por

$$r(t) = t \cdot q + (1 - t) \cdot p, \quad 0 \leq t \leq 1.$$

La velocidad en esta parametrización es $r'(t) = (-4, 2, 2)$, luego es constante en t y no depende del punto.

2. **Circunferencias.** En el espacio podemos considerar un círculo unidad C contenido en el plano $\{z = 3x - 2y\}$. Si asumimos que C proyecta en el plano (x, y) al círculo unidad centrado en el origen, una parametrización posible de C es

$$r : [0, 2\pi] \longrightarrow \mathbb{R}^3, \quad r(t) = (\cos t, \sin t, 3 \cos t - 2 \sin t), \quad 0 \leq t \leq 2\pi.$$

La velocidad en esta parametrización es $r'(t) = (-\sin t, \cos t, -3 \sin t - 2 \cos t)$.

3. **Hélices.** La imagen de la aplicación

$$r : [0, 2\pi] \longrightarrow \mathbb{R}^3, \quad r(t) = (\cos t, \sin t, t)$$

es una hélice que crece linealmente en la dirección vertical desde altura 0 a altura 2π . La velocidad de una partícula recorriendo la hélice con posiciones dadas por $r(t)$ es $r'(t) = (-\sin t, \cos t, 1)$.

3.3. Superficies

Una superficie es un objeto bidimensional. Esto es, un conjunto de puntos que es posible recorrer usando movimientos en dos direcciones. Las superficies se suelen ver en el 3-espacio. Si ignoramos el grosor, una hoja de papel, la hoja de un árbol o un mantel son ejemplos de superficies. Existen dos maneras de dar una superficie en el 3-espacio: mediante una parametrización o una ecuación.

3.3.1. Parametrización de una superficie

Una superficie parametrizada es la imagen de un subconjunto del plano vía una aplicación diferenciable al 3-espacio. Siendo precisos:

Una superficie en el espacio $\mathbb{R}^3$ de variables (x, y, z) es la imagen de una aplicación diferenciable

$$r : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \quad t \longmapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in \Omega.$$

La función r parametriza la superficie.

Ejemplos: 1. **Planos.** Sea $\Omega = \mathbb{R}^2$ y consideramos la parametrización

$$r(u, v) = (u, v, 2), \quad (u, v) \in \mathbb{R}^2.$$

La imagen de la aplicación r es un 2-plano π en $\mathbb{R}^3$. De hecho $\pi = \{z = 2\}$. Si consideramos

$$r(u, v) = (u, v, 3u - 2v + 1), \quad (u, v) \in \mathbb{R}^2,$$

también obtenemos un 2-plano en $\mathbb{R}^3$, distinto de π .

2. **Discos.** Sea $\Omega = \mathbb{D}^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$ un disco en el plano. La imagen de Ω por la parametrización $r(u, v) = (u, v, 0)$ es un disco contenido en el plano $\{z = 0\}$. En este caso la superficie no está curvada. Si en cambio consideramos la parametrización

$$s(u, v) = (u, v, \sqrt{1 - u^2 - v^2}), \quad (u, v) \in \Omega$$

la imagen es la hemisferio norte de la esfera unidad centrada en el origen $0 \in \mathbb{R}^3$. Esta superficie está curvada.

3. **Esferas.** Sea $\Omega = [0, 2\pi] \times [0, \pi] \subset \mathbb{R}^2$. La imagen de Ω por la parametrización

$$r(u, v) = (\cos u \sin v, \sin u \sin v, \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

está contenida en la esfera unidad $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Debido al hecho que existen elección de (u, v) de modo que todo punto de $\mathbb{S}^2$ es de la forma $r(u, v)$, la imagen es $\mathbb{S}^2$. Nótese que esta parametrización no es más que la expresión en coordenadas esféricas con la restricción del radio $\{r = 1\}$.

4. Integración

Ahora que las nociones básicas de geometría en el plano y el espacio han sido introducidas, definiremos operaciones sobre estos objetos. Estas operaciones nos permitirán calcular magnitudes como la longitud de una curva o el área de una superficie. A su vez, la integración da un fundamento riguroso a la noción del trabajo de una fuerza a lo largo de una curva o el flujo de un campo a través de una superficie.

4.1. Integrales en curvas

Las integrales sobre curvas también se llaman *integrales de línea*, que proviene del hecho que las curvas son objetos de dimensión uno.

4.1.1. Longitud de una curva plana

Definición 3. Sea C una curva plana descrita por $r : [a, b] \rightarrow \mathbb{R}^2$. La longitud $L(C)$ de la curva plana C es

$$L(C) = \int_a^b |r'(t)| dt.$$

Es sencillo comprobar que para otra parametrización $s : [c, d]$ con la misma curva imagen C se tiene

$$\int_c^d |s'(t)| dt = \int_a^b |r'(t)| dt.$$

Luego la cantidad $L(C)$ es independiente de la parametrización y sólo depende de C .

4.1.2. Integral de línea en el plano

La integral de línea de una función $f(x, y)$ en el plano sobre una curva C parametrizada por $r : [a, b] \rightarrow \mathbb{R}^2$ es

$$\int_C f = \int_a^b f(r(t)) \|r'(t)\| dt.$$

Esto generaliza la sección anterior, es decir, la longitud de una curva es la integral de línea de la función $f(x, y) = 1$ sobre la curva C que queremos medir:

$$L(C) = \int_C 1.$$

Es similar a la forma de calcular áreas y volúmenes de una región en el plano o en el espacio, que vimos anteriormente.

Además, el resultado es independiente de la parametrización r de la curva C y también de la orientación (es decir, si se recorre la curva al revés o no). Geométricamente, la gráfica de f es una superficie $z = f(x, y)$ en el espacio, por lo que la integral de línea se interpreta como el área de una valla (o cortina) entre la base de la imagen de C y la gráfica de f al igual que ocurre con las integrales de una variable.

Nota: Cuando una curva empieza y acaba en el mismo punto (es decir, es cerrada), la integral de línea tiene otra notación también:

$$\int_C f = \oint_C f.$$

Ejemplo 5. Sean $p, q \in \mathbb{R}^2$ dos puntos en el plano y $r(t) = t \cdot q + (1-t) \cdot p$, $0 \leq t \leq 1$ una parametrización del segmento C que les une. La longitud debería ser el módulo del vector $q - p$. En efecto,

$$L(C) = \int_0^1 |r'(t)| dt = \int_0^1 |q - p| dt = |q - p| \int_0^1 dt = |q - p|.$$

Ejemplo 6. Demostremos que la longitud de una circunferencia de radio r es $2\pi r$, es decir, busquemos hallar

$$\oint_C ds$$

siendo C la longitud de una circunferencia de radio r .

Por simplicidad, consideremos una circunferencia de radio r centrada en el origen, por lo que una posible parametrización es

$$\mathbf{r}(t) = (r \cos t, r \sin t) \quad 0 \leq t \leq 2\pi$$

Dado que

$$\begin{aligned} \mathbf{r}'(t) &= (-r \sin t, r \cos t) \\ \|\mathbf{r}'(t)\| &= \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} \\ &= \sqrt{r^2 (\sin^2 t + \cos^2 t)} \\ &= r \end{aligned}$$

Por lo tanto

$$\begin{aligned} L(C) &= \int_C 1 = \oint_C 1 \\ &= \int_0^{2\pi} r \, dt \\ &= 2\pi r \end{aligned}$$

Ejemplo 7. Sea $C \subset \mathbb{R}^2$ la elipse de semiejes $a, b > 0$ centrada en el origen con la parametrización

$$\mathbf{r}(t) = (a \cos t, b \sin t), \quad 0 \leq t \leq 2\pi.$$

Luego la longitud de esta elipse es

$$\begin{aligned} L(C) &= \int_0^{2\pi} \|\mathbf{r}'(t)\| \, dt = \int_0^{2\pi} \|(-a \sin t, b \cos t)\| \, dt = \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} \, dt = \\ &= 4 \int_0^{\pi/2} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} \, dt = 4a \int_0^{\pi/2} \sqrt{1 - \varepsilon \cos^2 t} \, dt \end{aligned}$$

donde $\varepsilon = \sqrt{1 - b^2/a^2}$ es la excentricidad de la elipse. La integral de la función $\sqrt{1 - \varepsilon \cos^2 t}$ no puede expresarse en términos de funciones elementales. La longitud de una elipse es una cantidad difícil de calcular.

4.1.3. Longitud de una curva en el espacio

El concepto de longitud para una curva en el espacio coincide con el de longitud de una curva plana. La única variación es que la imagen de $\mathbf{r}(t)$ está contenida en $\mathbb{R}^3$:

Definición 4. Sea $C \subset \mathbb{R}^3$ una curva espacial descrita por $\mathbf{r} : [a, b] \rightarrow \mathbb{R}^3$. La longitud $L(C)$ de la curva plana C es

$$L(C) = \int_a^b \|\mathbf{r}'(t)\| \, dt.$$

Dada otra parametrización $\mathbf{s} : [c, d]$ con la misma curva imagen C se tiene

$$\int_c^d \|\mathbf{s}'(t)\| \, dt = \int_a^b \|\mathbf{r}'(t)\| \, dt,$$

y la longitud $L(C)$ es independiente de la parametrización.

Ejemplo 8. Sea $C \subset \mathbb{R}^3$ la hélice descrita por

$$\mathbf{r} : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad \mathbf{r}(t) = (\cos t, \sin t, t).$$

La velocidad es $\mathbf{r}'(t) = (-\sin t, \cos t, 1)$. Luego la longitud es

$$L(C) = \int_0^{2\pi} \|(-\sin t, \cos t, 1)\| \, dt = \int_0^{2\pi} \left(\sqrt{\sin^2 t + \cos^2 t + 1} \right) \, dt = \int_0^{2\pi} \sqrt{2} \, dt = 2\sqrt{2}\pi.$$

4.1.4. Integral de línea en el espacio

Esta noción es el análogo 3-dimensional del caso plano. La integral de línea de una función $f(x, y, z)$ en $\mathbb{R}^3$ sobre una curva C parametrizada por $r : [a, b] \rightarrow \mathbb{R}^3$ es

$$\int_C f = \int_a^b f(r(t)) \|r'(t)\| dt.$$

Notad que no cambia nada en absoluto de la definición en comparación al caso plano. También en este caso, si la curva es cerrada se puede usar la notación $\oint$.

Ejemplo 9. Se desea evaluar la integral de línea

$$\int_C (x^2 + y^2 + z^2) ds$$

sobre la hélice

$$\begin{aligned} r &: [0, 2\pi] \rightarrow \mathbb{R}^3 \\ t &\mapsto (\cos t, \sin t, t) \end{aligned}$$

En primer lugar notemos que

$$r'(t) = (-\sin t, \cos t, 1)$$

por lo que

$$\begin{aligned} \|r'(t)\| &= \sqrt{(-\sin t)^2 + \cos^2 t + 1^2} \\ &= \sqrt{\sin^2 t + \cos^2 t + 1} \\ &= \sqrt{2} \end{aligned}$$

Y como

$$\begin{aligned} f(r(t)) &= \cos^2 t + \sin^2 t + t^2 \\ &= 1 + t^2 \end{aligned}$$

Entonces

$$\begin{aligned} \int_C (x^2 + y^2 + z^2) ds &= \int_0^{2\pi} (1 + t^2) \sqrt{2} dt \\ &= \sqrt{2} \left(t + \frac{t^3}{3} \right) \Big|_0^{2\pi} \\ &= \frac{2\sqrt{2}\pi(3 + 4\pi^2)}{3} \end{aligned}$$

4.2. Integrales de Superficie

Esta sección introducimos la integración de una función escalar a lo largo de una superficie. En el caso particular que la función escalar es constante igual a 1 obtendremos el área de una superficie.

4.2.1. Área de una superficie

La magnitud $\|v \times w\|$ del producto vectorial de dos vectores $v, w \in \mathbb{R}^3$ es el área del paralelogramo generado por v, w en el plano $\langle v, w \rangle$.

Definición 5. Sea S una superficie descrita por $r : \Omega \rightarrow \mathbb{R}^3$. El área $A(S)$ de la superficie S es

$$A(S) = \iint_{\Omega} \|\partial_u r \times \partial_v r\| du dv.$$

Además, a $n(u, v) = \partial_u r \times \partial_v r$ se le llama *vector normal*.

Ejemplo 10. Sea $\Omega = [0, 2\pi] \times [0, \pi] \subset \mathbb{R}^2$. La imagen de Ω por la parametrización

$$r(u, v) = (\rho \cos u \sin v, \rho \sin u \sin v, \rho \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

es la esfera $\mathbb{S}_\rho^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = \rho^2\}$ de radio ρ centrada en el origen. Además,

$$n(u, v) = \partial_u r \times \partial_v r = (-\rho^2 \cos u \sin^2 v, -\rho^2 \sin u \sin^2 v, -\rho^2 \cos v \sin v) = -\rho^2 \sin v \cdot (x, y, z).$$

El área de $\mathbb{S}_\rho^2$ es la integral del módulo del vector normal n a lo largo de Ω :

$$\begin{aligned} A(S) &= \iint_{\Omega} \|n\| du dv = \int_0^{2\pi} \int_0^\pi |-\rho^2 \sin v| du dv = \\ &= \rho^2 \int_0^{2\pi} \int_0^\pi |\sin v| du dv = 2\pi \rho^2 \int_0^\pi |\sin v| dv = 4\pi \rho^2. \end{aligned}$$

4.2.2. Integral de superficie

Definición 6. Sea S una superficie descrita por $r : \Omega \rightarrow \mathbb{R}^3$. La integral de f a lo largo de la superficie S es

$$\iint_{\Omega} f(r(u, v)) \|\partial_u r \times \partial_v r\| du dv.$$

Ejemplo 11. Consideramos $S = \mathbb{S}_\rho^2$ con la parametrización

$$r(u, v) = (\rho \cos u \sin v, \rho \sin u \sin v, \rho \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

y la función $f : \mathbb{S}_\rho^2 \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $f(x, y, z) = z$. Restringida a la superficie $\mathbb{S}_\rho^2$ esta función es

$$f|_S = (f \circ r)(u, v) = \rho \cos v.$$

Teniendo en cuenta que la función f tiene valores opuestos en el hemisferio norte y el hemisferio sur, la integral debería ser cero. El vector normal es

$$n = \partial_u r \times \partial_v r = (-\rho^2 \cos u \sin^2 v, -\rho^2 \sin u \sin^2 v, -\rho^2 \cos v \sin v) = -\rho^2 \sin v \cdot (x, y, z).$$

La integral de f a lo largo de S es

$$\begin{aligned} \iint_{\Omega} f \|n\| du dv &= \int_0^{2\pi} \int_0^\pi \rho \cos v |-\rho^2 \sin v| du dv = \\ &= \rho^2 \int_0^{2\pi} \int_0^\pi \rho \cos v |\sin v| du dv = 2\pi \rho^3 \int_0^\pi \cos v \sin v dv = 0. \end{aligned}$$

1. Campos vectoriales

Hasta ahora solo hemos visto integrales sobre curvas y superficies de *funciones escalares*, es decir, integrales de funciones que dan un número real (un *escalar*) a cada punto de la curva o de la superficie. Para dar el siguiente paso más allá necesitamos definir *campos vectoriales*.

Recordad que un vector ω en $\mathbb{R}^n$ sencillamente es una n -tupla de números, es decir, $\omega = (a_1, \dots, a_n) \in \mathbb{R}^n$. En particular, un vector de $\mathbb{R}^1$ es lo mismo que un único número, un escalar. Si las funciones de escalares asignan un número a cada punto, un campo de vectores es lo análogo pero con vectores en lugar de números. Dicho de otra forma, si nuestro dominio es el conjunto Ω (curve, superficie, etc.), las funciones escalares son

$$f : \Omega \rightarrow \mathbb{R} = \mathbb{R}^1,$$

y su generalización son los campos vectoriales

$$F : \Omega \rightarrow \mathbb{R}^n$$

Definición 1. Un *campo vectorial* en Ω (donde Ω puede ser una curva, superficie, $\mathbb{R}^n$, etc.) es una asignación de vectores a cada punto de Ω . Es decir,

$$F : \Omega \rightarrow \mathbb{R}^n,$$

donde $F(x) = \omega_x$ se llama el vector del campo F en el punto x .

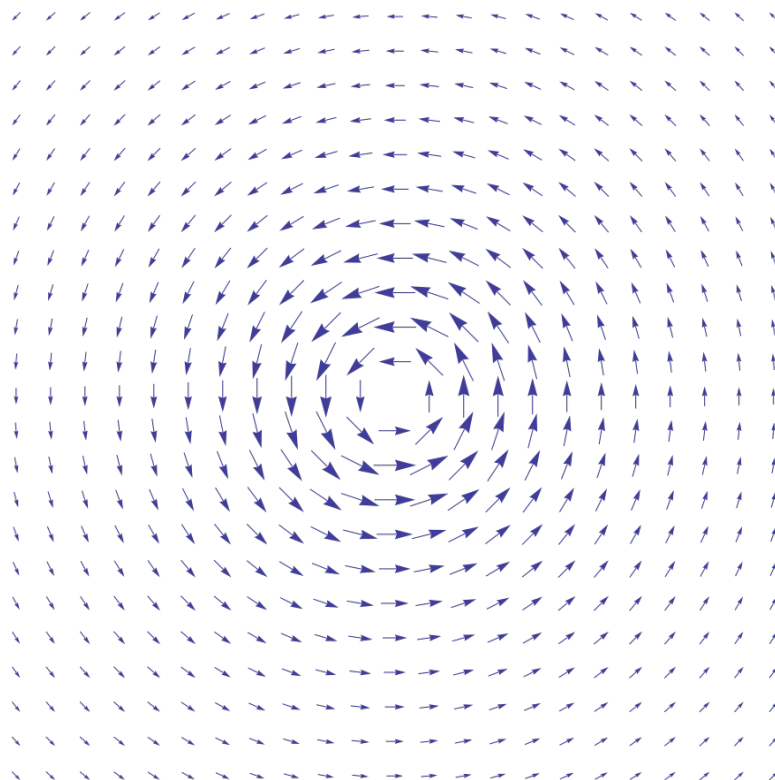
Dicho de otra forma, si nuestro dominio es el conjunto Ω (curve, superficie, etc.), las funciones escalares son

$$f : \Omega \rightarrow \mathbb{R} = \mathbb{R}^1,$$

y su generalización son los campos vectoriales

$$F : \Omega \rightarrow \mathbb{R}^n.$$

Típicamente, los campos vectoriales se representan gráficamente como un montón de flechas sobre el lugar sobre el que está el campo vectorial por razones obvias. Por ejemplo, un campo vectorial sobre un plano se representa así:



Ejemplo 1. Como en todo buen texto explicativo, ya hemos visto campos vectoriales antes pero no éramos conscientes de ello. Dada una función cualquiera $f : \mathbb{R}^n \rightarrow \mathbb{R}$, podemos tomar su *campo gradiente*,

$$\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

que acada punto $x \in \mathbb{R}^n$ le asigna el vector $\nabla f(x)$.

Una pregunta muy interesante es si todos los campos vectoriales son gradientes de alguna función. Los campos vectoriales que son gradientes de alguna función se llaman campos *conservativos* y, como podéis imaginar (porque tienen nombre) no todos los campos son conservativos. Veremos por qué se llaman así más adelante. Por ejemplo, pensad si la representación del campo vectorial anterior puede ser el gradiente de alguna función.

2. Integración

Anteriormente hemos visto cómo calcular integrales de funciones escalares tanto sobre curvas,

$$\int_C f = \int_a^b f(r(t)) \|r'(t)\| dt,$$

como sobre superficies,

$$\int_S f = \iint_{\Omega} f(r(u, v)) \|\partial_u r \times \partial_v r\| dudv.$$

En ambos casos necesitábamos una parametrización, tanto de la curva ($r : [a, b] \rightarrow C$) como de la superficie ($r : \Omega \rightarrow S$).

2.1. Sobre curvas

Definición 2. Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un campo vectorial y $C \subset \mathbb{R}^n$ una curva descrita por $r(t)$ con $a \leq t \leq b$. La integral de línea de F a lo largo de C es

$$\int_C F := \int_a^b \langle F(r(t)), r'(t) \rangle dt.$$

Ejemplo 2. 1. Calculemos la integral de línea del campo vectorial $F(x, y) = (y, x^2)$ en la curva C dada por $r(t) = (4 - t, 4t - t^2)$ con $0 \leq t \leq 3$. Según la descripción de la curva C : $x(t) = 4 - t$, $y(t) = 4t - t^2$. Luego

$$F(r(t)) = (4t - t^2, (4 - t)^2), \quad r'(t) = (-1, 4 - 2t), \quad F(r(t)) \cdot r'(t) = t^2 - 4t + (4 - 2t)(4 - t)^2.$$

La integral de línea de F a lo largo de C es

$$\int_C F = \int_0^3 t^2 - 4t + (4 - 2t)(4 - t)^2 dt = \frac{69}{2}.$$

Ejemplo 3. Calculemos la integral de línea del campo vectorial $F(x, y) = (x^3 - 2x^2, x - y/2)$ al movernos en la curva C descrita por $r(t) = (t, t^2)$ con $-1 \leq t \leq 2$. En este caso $r(t) = (x(t), y(t)) = (t, t^2)$ y tenemos

$$F(r(t)) = (t^3 - 2t^2, t - t^2/2), \quad r'(t) = (1, 2t) \implies F(r(t)) \cdot r'(t) = t^3 - 2t^2 + 2t^2 - t^3 = 0,$$

así que el vector tangente a la curva y el campo F son ortogonales. Luego la integral de línea de F a lo largo de C es

$$\int_C F = \int_{-1}^2 0 dt = 0$$

En general si el campo $F(r(t))$ es perpendicular al vector tangente $r'(t)$, i.e. $F(r(t)) \cdot r'(t) = 0$, la integral es cero.

2.2. Sobre superficies

Normalmente estas integrales las calcularemos de otra forma.

Definición 3. Sea $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vectorial y $S \subset \mathbb{R}^3$ una superficie parametrizada por $r : \Omega \rightarrow S$. La integral de superficie de F en S es

$$\int_S F = \iint_{\Omega} \langle F(r(u, v)), (\partial_u r \times \partial_v r) \rangle du dv.$$

Puesto que $\langle a, b \times c \rangle$ es igual al determinante

$$\begin{vmatrix} \cdots & a & \cdots \\ \cdots & b & \cdots \\ \cdots & c & \cdots \end{vmatrix},$$

el cálculo de estas integrales es muy sencillo.

3. Teoremas

3.1. Operadores

Antes de hablar de los teoremas relacionados con integración de campos vectoriales, es conveniente dar la definición de los cuantos *operadores* (que no son más que *operaciones* que se hacen sobre funciones).

Operadores que actúan sobre campos escalares (i.e., funciones escalares).

■ El operador gradiente

El gradiente de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es $\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$. En consecuencia, el operador gradiente se define como

$$\nabla := \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right). \quad (1)$$

Puedes saber más sobre este operador [aquí](#).

■ El operador laplaciano

El laplaciano de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es $\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$. En consecuencia, el operador laplaciano se define como

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}, \text{ o también } \Delta = \nabla^2. \quad (2)$$

Observa que este operador es de segundo orden, es decir intervienen las derivadas segundas para definirlo. Puedes saber más sobre este operador [aquí](#).

Operadores que actúan sobre campos vectoriales.

Si Ω es un recinto en $\mathbb{R}^3$, y $F : \Omega \rightarrow \mathbb{R}^3$, $F = (F_1, F_2, F_3)$, es un campo escalar en Ω , entonces definimos los siguientes operadores diferenciales.

■ El operador divergencia

La divergencia de F es el campo escalar $\text{Div}(F) = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}$. En consecuencia, el operador divergencia se define como:

$$\text{Div} := \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}, \text{ o también } \text{Div}(F) = \nabla \cdot F \quad (3)$$

Puedes saber más sobre este operador [aquí](#).

■ El operador rotacional

El rotacional de F es el campo vectorial

$$\text{rot}(F) = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \vec{i} + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \vec{j} + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \vec{k}.$$

En consecuencia, el operador rotacional se define como:

$$\text{rot}(F) = \nabla \times F = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{vmatrix}. \quad (4)$$

Puedes saber más sobre este operador [aquí](#).

3.2. Teorema de Green

El Teorema de Green aplica sólo a integrales de línea a lo largo de una **curva plana cerrada**. Sea F un campo vectorial en el plano $\mathbb{R}^2$. Expresamos este campo como

$$F(x, y) = (P(x, y), Q(x, y))$$

Sea $C \subset \mathbb{R}^2$ una curva plana cerrada y orientada. Esta curva es 1-dimensional y la integral de línea de F a lo largo de C es equivalente al cálculo de una integral en 1 variable. El dominio Ω es definido como la región acotada que queda encerrada por C . En otras palabras, la frontera de Ω es C .

Teorema 3.1. (*Teorema de Green*) Sea Ω una dominio en el plano con frontera una curva conexa orientada cerrada C . Dado un campo vectorial $F(x, y) = (P(x, y), Q(x, y))$,

$$\int_C F = \iint_{\Omega} (\partial_x Q - \partial_y P) \, dxdy$$

Observación: Para aplicar el teorema **no** es necesario tener la curva C descrita por $r(t)$. Sólo es necesario describir el dominio interior Ω .

Ejemplo 4. Circunferencias. Sea C la circunferencia de radio ρ centrada en el origen. El dominio Ω es el disco de radio ρ centrado en el origen: $\Omega = \{(r, \theta) : r \leq \rho\} = \{(x, y) : x^2 + y^2 \leq \rho^2\}$. Consideramos $F(x, y) = (P(x, y), Q(x, y)) = (y^2 + y + e^x, 2xy - e^{2y} + \tan y)$. La integral de línea de F a lo largo de C es

$$\int_C F \cdot dl = \iint_{\Omega} Q_x - P_y \, dxdy = \iint_{\Omega} -1 \cdot dxdy = -\text{Area}(\Omega) = -\pi\rho^2.$$

Ejemplo 5. Sea $C \subset \mathbb{R}^2$ el triángulo de vértices $(-1, 0), (1, 0), (0, 1)$ y $F(x, y) = (e^y, -\sin(\pi x))$. El dominio Ω es el interior del triángulo, cuyos límites de integración se pueden tomar como $0 \leq y \leq 1$ y $y - 1 \leq x \leq 1 - y$. Vía el teorema de Green, la integral de línea de F a lo largo de C se puede calcular integrando $Q_x - P_y$ en el interior del triángulo:

$$I = \int_{\Omega} (Q_x - P_y) \cdot dxdy = \int_0^1 \int_{y-1}^{1-y} (-\pi \cos(\pi x) - e^y) \cdot dxdy = -2(e - 2) - \frac{4}{\pi}.$$

Ejemplo 6. Sea $C \subset \mathbb{R}^2$ una curva plana cerrada cualquiera y $F = (P, Q)$ una fuerza con $Q_x = P_y$. Entonces

$$\int_C F \cdot dl = 0.$$

Instancias de este hecho son las fuerzas

$$F(x, y) = (x, 0), F(x, y) = (0, y), F(x, y) = (x, y) \text{ y } F(x, y) = (y, x).$$

El ejemplo anterior motiva volver a discutir sobre *campos conservativos*.

Definición 4. Un campo vectorial F es conservativo si es igual al gradiente de alguna función, $\nabla f = F$.

La propiedad fundamental de los campos conservativos es que **su integral sobre un camino cerrado es siempre cero**. Esto es muy fácil de ver con el teorema de Green, hagámoslo.

Asumamos que F es un campo conservativo, e intentemos calcular

$$\int_C F$$

donde C es una curva cerrada en el plano. Como F es conservativo, $F = \nabla f$ para alguna función f (que en general desconocemos). Así pues,

$$\int_C F = \int_C \nabla f.$$

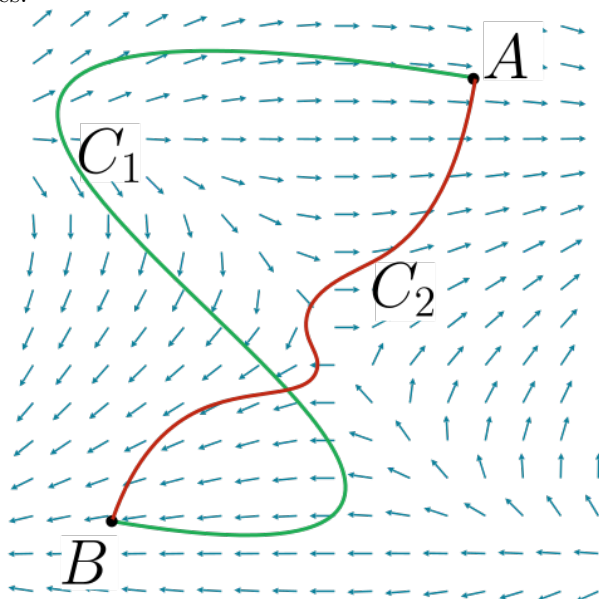
Ahora podemos aplicar el teorema de Green, que nos dice que debemos integrar sobre la región R que queda encerrada por la curva C . Además, como $\nabla f = (\partial_x f, \partial_y f)$, el teorema de Green nos dice que

$$\int_C \nabla f = \int_R (\partial_x(\partial_y f) - \partial_y(\partial_x f)).$$

Por otro lado, dadas las circunstancias adecuadas de diferenciación, sabemos que la derivada cruzada de f es la misma sea cual sea el orden, es decir: $\partial_x \partial_y f = \partial_y \partial_x f$. Por lo tanto,

$$\int_C \nabla f = \int_R (\partial_x(\partial_y f) - \partial_y(\partial_x f)) = \int_R 0 = 0.$$

Esto además nos permite probar que si F es un campo conservativo y calculamos la integral de línea de una curva C que va entre los puntos A y B en el plano, el resultado solo depende de los puntos A y B . Es decir, la integral sobre cualquier curva que va del punto A al punto B es la misma. Esto se puede probar encontrando regiones cerradas por dos curvas C_1 y C_2 y aplicando el resultado anterior repetidas veces.



3.3. Teorema de Stokes

Si nos fijamos, lo que aparece en el teorema de Green se parece sospechosamente a lo que obtenemos al hacer el rotacional de un campo en $\mathbb{R}^3$. No es lo mismo, porque el teorema de Green solo habla de campos vectoriales en el plano, y el rotacional solo es para campos vectoriales en $\mathbb{R}^3$. Esto nos lleva al teorema de Stokes, que es la generalización del teorema de Green en 3 dimensiones.

Teorema 3.2. (Teorema de Stokes) Sea $\mathcal{S}$ una superficie de $\mathbb{R}^3$ cuya frontera $\partial\mathcal{S}$ es una unión finita de curvas cerradas simples, orientada positivamente respecto de la normal exterior a $\mathcal{S}$. Sea $F = (F_1, F_2, F_3)$ un campo definido en cada punto de $\mathcal{S}$. Entonces

$$\oint_{\partial\mathcal{S}} F = \iint_{\mathcal{S}} \text{rot}(F). \quad (5)$$

Ejemplo 7. Calcula la integral del campo $F = (\sin(x), e^y + x^2, z^4 + 2x^2)$ sobre la frontera del triángulo de vértices $A = (3, 0, 0)$, $B = (0, 2, 0)$ y $C = (0, 0, 1)$.

Para calcular esta integral de línea basta usar el Teorema de Stokes y por esto calcular

$$\iint_{\mathcal{S}} \text{rot}(F) \cdot d\vec{S}.$$

En concreto, como $F_1 = \sin(x)$, $F_2 = x^2 + e^y$, $F_3 = z^4 + 2x^2$, su rotacional es

$$\text{rot}(F) = \nabla \times F = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial_x & \partial_y & \partial_z \\ \sin(x) & x^2 + e^y & 2x^2 + z^4 \end{vmatrix} = (-4x)\vec{j} + (2x)\vec{k}.$$

Para calcular la integral de superficie tenemos que parametrizar la superficie $\mathcal{S}$ de alguna forma. En este caso, como el triángulo está contenido en el plano $-2x - 3y + 3z + 6 = 0$ tenemos que la aplicación de

parametrización G está definida sobre el triángulo plano Ω de vértices $P = (0, 0)$, $Q = (3, 0)$ y $R = (0, 2)$ por la fórmula:

$$G(x, y) = \left(x, y, \frac{-6 + 2x + 3y}{3} \right), \quad (6)$$

para (x, y) en el triángulo

$$\Omega = \{(x, y) | 0 \leq x \leq 3, 0 \leq y \leq -\frac{2}{3}x + 2\}.$$

El vector normal asociado a esta parametrización se calcula como sigue. En primer lugar calculamos las derivadas parciales de la parametrización:

$$\partial_x G = (1, 0, 2/3), \quad \partial_y G = (0, 1, 1), \quad (7)$$

lo que da el campo normal a la superficie

$$n(x, y) = \partial_x G \times \partial_y G = (-2/3, -1, 1). \quad (8)$$

Observa que está orientado positivamente porque su componente z , $z = 1$, es siempre positiva. Con esta orientación se tiene que

$$\iint_S \text{rot}(F) \cdot d\vec{S} = \int_{\Omega} (0, -4x, 2x) \cdot \left(-\frac{2}{3}, -1, 1 \right) dx dy = \int_{\Omega} (6x) dx dy.$$

Calculemos esta integral:

$$\int_{\Omega} (6x) dx dy = 6 \int_0^3 \int_0^{-\frac{2}{3}x+2} x dy dx = 6 \int_0^3 x \left(-\frac{2}{3}x + 2 \right) dx = 18.$$

En consecuencia, el valor de la integral de línea sobre el triángulo que se pedía inicialmente es 18.

3.4. Teorema de la divergencia

El teorema de la divergencia, o teorema de Gauss, es como el teorema de Stokes pero si añadimos una dimensión más a los dominios de integración. Es decir, intervienen un sólido en tres dimensiones y su frontera (que es una superficie), en comparación al caso del teorema de Stokes donde intervienen una superficie y su frontera (que es una curva).

Teorema 3.3. (*Teorema de la divergencia*) Sea S una superficie cerrada que limita una región $\mathcal{W}$ en $\mathbb{R}^3$. Supongamos que S es suave a trozos y orientada por vectores normales que apuntan al exterior de $\mathcal{W}$. Sea F un campo vectorial cuyo dominio contiene a $\mathcal{W}$. Entonces

$$\iint_S F = \iiint_{\mathcal{W}} \text{div}(F). \quad (9)$$

Ejemplo 8. Calcula la integral $\iint_S F$ donde S es la frontera de la caja $\mathcal{W} = [0, 2] \times [0, 3] \times [0, 1]$, y el campo viene dado por $F = (x^2, z^4, e^z)$.

En primer lugar calcularemos la divergencia del campo dado. En concreto

$$\text{div}(F) = \frac{\partial x^2}{\partial x} + \frac{\partial z^4}{\partial y} + \frac{\partial e^z}{\partial z} = 2x + e^z.$$

Por otro lado, el sólido $\mathcal{W}$ ya tiene una descripción adecuada para integrar. En consecuencia,

$$\begin{aligned} \iiint_{\mathcal{W}} \text{div}(F) &= \int_0^2 \left(\int_0^3 \left(\int_0^1 (2x + e^z) dz \right) dy \right) dx = \\ &= \int_0^2 \left(\int_0^3 (2x + e - 1) dy \right) dx = 6e + 6. \end{aligned}$$

Luego el flujo pedido es $6e + 6$ en la dirección de la normal exterior a la frontera de la caja.

Nota: Como se puede observar, los teoremas de Green, Stokes y de la divergencia tienen la misma filosofía, la integral sobre la frontera de algo es otra integral sobre su interior. Estos tres teoremas son casos particulares de una teorema más general, que se suele llamar *Teorema de Stokes generalizado* o, más dramáticamente, *teorema fundamental del cálculo multivariable*. Como podéis imaginar, el teorema fundamental del cálculo (el que relaciona integrales con sus derivadas) tiene mucho que ver, y es también un caso particular. Las matemáticas involucradas son técnicas para lo que concierne a este curso.

1. Integración en varias variables

El cálculo de una integral múltiple (en varias variables) se reduce a ir calculando integrales de una variable en el orden especificado. Intuitivamente, al igual que en el caso de una variable, las integrales múltiples calculan volúmenes. Veamos un ejemplo en el caso de dos variables:

Ejemplo 1. Supongamos que queremos calcular la integral múltiple

$$I = \int_0^1 \int_2^3 (6x + 6y^2) dx dy$$

donde las variables x, y tienen de límites $[2, 3]$ y $[0, 1]$ respectivamente. Para realizar la integración, primero empezamos integrando la variable x , manteniendo y constante,

$$I = \int_0^1 (3x^2 + 6y^2 x) \Big|_{x=2}^{x=3} dy,$$

y luego integramos sobre y ,

$$I = \int_0^1 (15 + 6y^2) dy = (15y + 2x^3) \Big|_{y=0}^{y=1} = 17.$$

En la integral anterior se puede comprobar que el orden de integración no importa, es decir,

$$\int_2^3 \int_0^1 (6x + 6y^2) dx dy = \int_0^1 \int_2^3 (6x + 6y^2) dy dx.$$

Sin embargo, observa que el límite de integración hay que ordenarlo también. Esto sucede gracias al teorema de *Fubini*, que no vamos a enunciar porque tiene unas hipótesis técnicas. En la práctica, cualquier función razonable cumple el teorema de Fubini.

Como decíamos anteriormente, la integral $\int_c^d \int_a^b f(x, y) dx dy$ es el volumen que la gráfica de $f(x, y)$ deja sobre el rectángulo $[a, b] \times [c, d]$ (si $|f| > 0$ en ese rectángulo). En general no haremos integrales sobre rectángulos, sino sobre regiones R del plano (o de $\mathbb{R}^3$). En ese caso, se puede escribir

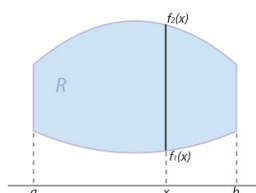
$$\iint_R f(x, y) dx dy.$$

De esta forma, cuando $f(x, y) = 1$, obtenemos:

$$\iint_R 1 dx dy = \text{Volumen del sólido que tiene base } R \text{ y altura } 1 = \text{Área de } R.$$

A veces, las regiones están descritas con funciones en las variables x o y :

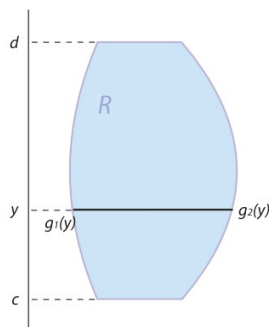
- Sugongamos una región R definida por las funciones $f_1(x)$ y $f_2(x)$, como por ejemplo



Debemos entonces integrar primero en y para tener en cuenta todas las secciones verticales, es decir

$$\iint_R F = \int_a^b \int_{f_1(x)}^{f_2(x)} F(x, y) dy dx$$

- Sugongamos una región R definida por las funciones $g_1(y)$ y $g_2(y)$, como por ejemplo



Debemos entonces integrar primero en x para tener en cuenta todas las secciones horizontales, es decir

$$\iint_R F = \int_c^d \int_{g_1(y)}^{g_2(y)} F(x, y) dx dy$$

A veces puede ocurrir también que las dos estrategias anteriores no funcionen, y tengamos que separar la región en trozos donde sí funcionen las estrategias anteriores. Esto es lo mismo que ocurre con las funciones definidas a trozos en una variable: hay que dividir la integral según los trozos y luego sumar todos los resultados.

Por último, como es evidente, se puede integrar en cuantas variables se quiera. Por ejemplo, se puede integrar en 3 variables, y el proceso es exactamente igual al anterior. En particular, se tiene que para la función $f(x, y, z) = 1$ y el sólido V ,

$$\int \int \int_V 1 dx dy dz = \text{Volumen del sólido } V.$$

De esta forma, si se sabe que un objeto tridimensional V tiene función de densidad $\rho(x, y, z)$ que depende del punto (x, y, z) donde se mida, se tiene que

$$\int \int \int_V \rho dx dy dz = \text{Masa del sólido } V,$$

que intuitivamente se entiende como masa=densidad·volumen.

En cualquier número de variables, si la función a integrar $f(x, y)$ es un producto de funciones $g(x)h(y)$, se puede ver fácilmente que

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

Nota: Es común que, cuando el contexto está claro, se eviten varios símbolos de la integral. Por ejemplo, si está claro por el contexto que estamos en el plano, todas las siguientes notaciones pueden aparecer en un libro: $\int \int_R f(x, y) dx dy$, $\int_R f(x, y) dx dy$, $\int_R f$.

Vamos a motivar la siguiente sección dando un ejemplo, que es muy difícil de calcular sin un *cambio de variable*.

Ejemplo 2. Calculemos el área del círculo de radio ℓ y centrado en el origen:

$$R = \{x^2 + y^2 \leq \ell^2\}.$$

En primer lugar, para calcular un área de una región sencillamente tenemos que calcular

$$A = \int \int_R 1 dx dy.$$

Reexpresamos R ,

$$R = \{-\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}\},$$

de tal forma que queda como una de las regiones que escribíamos anteriormente:

$$A = \int_{-\ell}^{\ell} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy dx,$$

donde también sabemos que x varía en $[-\ell, \ell]$. Hacemos la primera integral iterada:

$$A = \int_{-\ell}^{\ell} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy dx = \int_{-\ell}^{\ell} 2\sqrt{1-x^2} dx.$$

Esta integral es complicada. Hay que saber que este tipo de integrales se resuelven con el cambio de variable $x = \ell \sin(u)$, $dx = \ell \cos(u) du$,

$$A = \int_{3\pi/2}^{\pi/2} \ell^2 \cos^2(u) du.$$

Además hay que usar el truco de tomar $\cos^2(u) = \cos(2u)/2 + 1/2$:

$$A = \int_{3\pi/2}^{\pi/2} \left(\ell^2 \frac{\cos(2u)}{2} + \ell^2 \frac{1}{2} \right) du = \ell^2 \left(\frac{\sin(2u)}{4} + \frac{1}{2} u \right) \Big|_{u=3\pi/2}^{u=\pi/2} = \ell^2 \pi.$$

La integral, que ya ha necesitado un cambio de variable al llegar a la integración en solo una variable, sería muchísimo más simple cambiando de variables en $\mathbb{R}^2$ a coordenadas polares.

2. Cambio de variable

Por incontables razones (al menos dos: simplificar la integral y simplificar la región de integración), es conveniente a veces hacer un cambio de variable, como ejemplificábamos antes. En analogía con el caso de una variable, se tiene el siguiente resultado, que lo enunciaremos para 2 variables pero vale para cualquier número de variables.

Proposición 1. Dada una función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ integrable y una región R en $\mathbb{R}^2$, si $\phi(u, v) = (x(u, v), y(u, v))$ es un cambio de variable $(u, v) \mapsto (x, y)$ (es decir, una aplicación biectiva), entonces

$$\int \int_R f(u, v) du dv = \int \int_R f(x, y) |det D\phi|,$$

donde $D\phi$ es la matriz diferencial y a la cual se tiene que hacer el valor absoluto del determinante.

Veamos el ejemplo anterior:

Ejemplo 3. Calculemos el área del círculo de radio ℓ y centrado en el origen:

$$R = \{x^2 + y^2 \leq \ell^2\}.$$

Podemos cambiar a coordenadas polares, que suelen funcionar mejor cuando tenemos cosas redondas o aparecen senos y cosenos en la integral. Como

$$\begin{aligned} x &= r \cos(\theta) \\ y &= r \sin(\theta), \end{aligned}$$

esto nos define una función de cambio de variable $\phi(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$. Como anteriormente, calculamos

$$A = \int \int_R 1 dx dy.$$

Como queremos cambiar de variables, tenemos que calcular la diferencial de ϕ y su determinante:

$$\det D\phi = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r.$$

Tendríamos que tomar también el valor absoluto, pero r ya toma valores positivos, así que es igual a su valor absoluto. Al hacer el cambio de variable, θ varía en $[0, 2\pi]$ (para cubrir todos los ángulos del círculo) y r varía en $[0, \ell]$. Queda

$$A = \int_0^{2\pi} \int_0^{\ell} r dr d\theta = \int_0^{2\pi} \frac{r^2}{2} \Big|_{r=0}^{r=\ell} d\theta = \int_0^{2\pi} \frac{\ell^2}{2} d\theta = \frac{\ell^2}{2} \theta \Big|_{\theta=0}^{\theta=2\pi} = \pi \ell^2.$$

Este cálculo ha sido mucho más sencillo.

Ahora damos los cambios típicos de coordenadas.

2.1. A polares

Como el cambio de variable está definido por $\phi(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$, su diferencial es

$$D\phi = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}.$$

A su vez, $|\det D\phi| = r$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y) dx dy = \int_R f(r \cos \theta, r \sin \theta) r dr d\theta.$$

2.2. A cilíndricas

El cambio a cilíndricas es como el de polares con, además, una z : $\phi(r, \theta, z) = (r \cos \theta, r \sin \theta, z) = (x, y, z)$

$$D\phi = \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A su vez, $|\det D\phi| = r$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y, z) dx dy dz = \int_R f(r \cos \theta, r \sin \theta, z) r dr d\theta dz.$$

2.3. A esféricas

Como el cambio de variable está definido por $\phi(r, \phi, \theta) = (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) = (x, y, z)$, su diferencial es

$$D\phi = \begin{pmatrix} \cos \phi \sin \theta & -r \sin \phi \sin \theta & r \cos \phi \cos \theta \\ \sin \phi \sin \theta & r \cos \phi \sin \theta & r \sin \phi \cos \theta \\ \cos \theta & 0 & -r \sin \theta \end{pmatrix}.$$

A su vez, $|\det D\phi| = r^2 \sin \theta$, donde ya no aparece el valor absoluto porque r es siempre positivo. Así:

$$\int_R f(x, y, z) dx dy dz = \int_R f(r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta) r^2 \sin \theta dr d\phi d\theta.$$

Ejemplo 4. Esto nos permite calcular el volumen de una esfera de radio ℓ ,

$$V = \int_R 1 dx dy dz,$$

donde R se puede describir como $\{x^2 + y^2 + z^2 \leq \ell^2\}$. Es mucho más sencillo cambiar a coordenadas esféricas, con las variables los valores $r \in [0, \ell]$, $\phi \in [0, 2\pi]$ y $\theta = [0, \pi]$:

$$V = \int_0^\pi \int_0^{2\pi} \int_0^\ell r^2 \sin \theta dr d\phi d\theta = \int_0^{2\pi} d\phi \int_0^\ell r^2 dr \int_0^\pi \sin \theta d\theta,$$

porque la función que nos queda es un producto de funciones en cada variable. Cuando se calcula esto último da

$$V = \frac{4\pi}{3} \ell^3.$$

Nota: Existen unas [coordenadas esféricas para las esferas de cualquier dimensión](#), muy parecidas a las que hemos visto nosotros para 3 dimensiones. Eso permite calcular el [volumen de las bolas de \$n\$ dimensiones](#).

3. Repaso geométrico

En esta primera sección recordamos los conceptos geométricos básicos. A partir de ellos podremos definir nociones como la longitud y el trabajo de una fuerza.

3.1. Operaciones con vectores

Las dos operaciones relevantes para el cálculo multivariable son:

- (1) La proyección de un vector sobre otro, en el plano o el espacio.
- (2) Dados dos vectores en el espacio, construir un tercer vector ortogonal.

La primera operación es el producto escalar y la segunda es el producto vectorial.

Producto escalar

Dados dos vectores $v, w \in \mathbb{R}^2$ de coordenadas $v = (v_1, v_2)$, $w = (w_1, w_2)$. La proyección de w sobre v es el vector de dirección v y módulo

$$w \cdot v := v_1 w_1 + v_2 w_2 \in \mathbb{R}.$$

En el caso de tener vectores en el espacio $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3) \in \mathbb{R}^3$, la proyección de w sobre v es el vector que tiene dirección v y módulo

$$w \cdot v := v_1 w_1 + v_2 w_2 + v_3 w_3 \in \mathbb{R}.$$

En ambos casos, el **número real** $w \cdot v$ se conoce como el producto escalar de v y w .

Dos vectores v, w son ortogonales si $w \cdot v = 0$.

Producto vectorial

Esta operación es exclusiva de vectores en el espacio $\mathbb{R}^3$. Dados dos vector $v, w \in \mathbb{R}^3$, el producto vectorial de v y w es el **vector** ortogonal a v y w de módulo el área con signo del paralelogramo generado por v y w .

En coordenadas, $v = (v_1, v_2, v_3)$, $w = (w_1, w_2, w_3)$, el producto vectorial de v y w es

$$\begin{aligned} v \times w &:= \det \begin{pmatrix} v_1 & w_1 & \vec{i} \\ v_2 & w_2 & \vec{j} \\ v_3 & w_3 & \vec{k} \end{pmatrix} = (v_2 w_3 - v_3 w_2) \cdot \vec{i} + (v_3 w_1 - v_1 w_3) \cdot \vec{j} + (v_1 w_2 - v_2 w_1) \cdot \vec{k} = \\ &= (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1). \end{aligned}$$

Notemos que $v \cdot (v \times w) = 0$ y $w \cdot (v \times w) = 0$.

3.2. Curvas

En una primera aproximación, una curva es un objeto unidimensional. Es decir, un conjunto de puntos de cierto espacio que cuando se recorre sólo se tiene la libertad de moverse en una dirección, con dos posibles sentidos. Un vehículo desplazándose a lo largo de una calle de un solo carril está restringido a ir hacia delante o hacia atrás: la calle puede curvarse pero sigue habiendo sólo una dirección. Un funambulista al que no se le permita saltar está obligado a moverse a lo largo de un objeto unidimensional, el cable donde se apoya. La calle por la que se desplaza el vehículo se dibuja mediante una curva en un mapa de la ciudad correspondiente, es lo que denominaremos una curva plana. El cable del funambulista vive dentro del 3-espacio y dibuja una curva en el espacio.

3.2.1. Curvas en el plano

Sea $I \subset \mathbb{R}$ un intervalo en la recta $\mathbb{R}$ de variable independiente t . Una curva en el plano $\mathbb{R}^2$ de variables (x, y) es la imagen de una aplicación diferenciable

$$r : I \subset \mathbb{R} \longrightarrow \mathbb{R}^2, \quad t \longmapsto r(t) = (x(t), y(t)).$$

Se dice que la función r parametriza la curva.

Ejemplos: 1. **Segmentos.** Sea $I = [0, 1]$ y $r(t) = (t, 0)$. La curva imagen es el segmento en el plano $\mathbb{R}^2$ que une los puntos $(0, 0)$ y $(1, 0)$. Si $I = [-1, 0]$ y $r(t) = (t, t + 1)$. La curva es el segmento que une los puntos $(-1, 0)$ y $(0, 1)$. En general es segmento que une los punto $p, q \in \mathbb{R}^2$ se puede parametrizar con

$$r(t) = t \cdot q + (1 - t) \cdot p, \quad 0 \leq t \leq 1.$$

2. **Circunferencias.** Sea $I = [0, 2\pi]$ y $r(t) = (\cos t, \sin t)$. La curva es la circunferencia de radio 1 centrada en el origen $(0, 0)$. Para describir la circunferencia de radio a centrada en el origen podemos usar $I = [0, 2\pi]$ y $r(t) = (a \cos t, a \sin t)$.

3. Sea $I = [0, \pi]$ y $r(t) = (\cos t, \sin t)$. Entonces la curva es la parte de la circunferencia unidad centrada en el origen contenida en el semiplano superior $\{(x, y) : y \geq 0\}$.

4. **Elipses.** Sea C la elipse de semiejes $a, b \geq 0$:

$$C = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\}.$$

Una posible parametrización es $r(t) = (a \cos t, b \sin t)$, $0 \leq t \leq 2\pi$.

La función $r(t)$ suele interpretarse como la posición de una partícula moviéndose a lo largo del tiempo t . Esto motiva la siguiente

Definición 1. La velocidad de una curva $r : I \rightarrow \mathbb{R}^2$ es el vector $r'(t) = (x'(t), y'(t))$.

Ejemplos: 1. **Segmentos.** Sea $r(t) = (t, 0)$. La velocidad es $r'(t) = (1, 0)$. Si la posición está descrita por la curva $r(t) = (t^2 + t, te^t)$, la velocidad es $r'(t) = (2t + 1, e^t + te^t)$. Para un segmento general $r(t) = t \cdot q + (1 - t) \cdot p$ la velocidad es constante igual a $r'(t) = q - p$.

2. **Circunferencias.** En el caso de la circunferencia unidad $r(t) = (\cos t, \sin t)$ la velocidad es $r'(t) = (-\sin t, \cos t)$. Si se tiene la circunferencia de radio a descrita por $r(t) = (a \cos t, a \sin t)$, la velocidad es $r'(t) = (-a \sin t, a \cos t)$.

3. **Elipses.** Sea C la elipse parametrizada por

$$r(t) = (a \cos t, b \sin t), \quad 0 \leq t \leq 2\pi.$$

La velocidad es $r'(t) = (-a \sin t, b \cos t)$.

En ciertos casos es posible dar una curva $C \subset \mathbb{R}^2$ con una ecuación implícita, i.e. $C = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$. Esa caracterización será de poco uso en los procesos de integración pero es útil para el cálculo de la recta tangente.

Ejemplos: 1. **Rectas.** Consideremos la recta $l \subset \mathbb{R}^2$ con vector director $v = (a, b)$. Supongamos $0 \in l$. Entonces una ecuación implícita para la recta es $f(x, y) = -bx + ay$, es decir, $l = \{(x, y) : -bx + ay = 0\}$.

2. **Circunferencias.** Sea $C \subset \mathbb{R}^2$ la circunferencia unidad centrada en el origen. Una ecuación implícita viene dada por $f(x, y) = x^2 + y^2 - 1$:

$$C = \{(x, y) : \|(x, y)\|^2 = 1\} = \{(x, y) : x^2 + y^2 = 1\} = \{(x, y) : x^2 + y^2 - 1 = 0\}.$$

3.2.2. Curvas en el espacio

La noción de curva sigue siendo la imagen de un intervalo $I \subset \mathbb{R}$ por una aplicación continua r . Para curvas planas la aplicación es $r : I \rightarrow \mathbb{R}^2$. Para una curva en el espacio sólo tenemos que admitir que el conjunto de llegada sea $\mathbb{R}^3$.

Una curva en el espacio $\mathbb{R}^3$ de variables (x, y, z) es la imagen de una aplicación diferenciable

$$r : I \subset \mathbb{R} \rightarrow \mathbb{R}^3, \quad t \mapsto r(t) = (x(t), y(t), z(t)).$$

Decimos que la función r parametriza la curva. La definición es análoga al caso de curvas plana añadiendo una tercera componente $z(t)$.

La noción de velocidad es análogo que en el caso de curvas planas:

Definición 2. La velocidad de una curva $r : I \rightarrow \mathbb{R}^3$ es el vector $r'(t) = (x'(t), y'(t), z'(t))$.

Siendo más precisos, la velocidad de una curva parametrizada por $r(t) = (x(t), y(t), z(t))$ en el punto $r(t_0)$ es el vector $r'(t_0) = (x'(t_0), y'(t_0), z'(t_0))$.

Ejemplos: 1. **Segmentos.** Consideremos los puntos $(1, 0, 2), (-3, 2, 4) \in \mathbb{R}^3$. El segmento que los une es una curva C parametrizada por

$$r : [0, 1] \rightarrow \mathbb{R}^3, \quad r(t) = t \cdot (-3, 2, 4) + (1 - t) \cdot (1, 0, 2) = (-4t + 1, 2t, 2t + 2), \quad 0 \leq t \leq 1.$$

Observemos que efectivamente $r(0) = (1, 0, 2)$ y $r(1) = (-3, 2, 4)$. En general dados dos puntos $p, q \in \mathbb{R}^3$ en el espacio, el segmento que los une se parametriza por

$$r(t) = t \cdot q + (1 - t) \cdot p, \quad 0 \leq t \leq 1.$$

La velocidad en esta parametrización es $r'(t) = (-4, 2, 2)$, luego es constante en t y no depende del punto.

2. **Circunferencias.** En el espacio podemos considerar un círculo unidad C contenido en el plano $\{z = 3x - 2y\}$. Si asumimos que C proyecta en el plano (x, y) al círculo unidad centrado en el origen, una parametrización posible de C es

$$r : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad r(t) = (\cos t, \sin t, 3 \cos t - 2 \sin t), \quad 0 \leq t \leq 2\pi.$$

La velocidad en esta parametrización es $r'(t) = (-\sin t, \cos t, -3 \sin t - 2 \cos t)$.

3. **Hélices.** La imagen de la aplicación

$$r : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad r(t) = (\cos t, \sin t, t)$$

es una hélice que crece linealmente en la dirección vertical desde altura 0 a altura 2π . La velocidad de una partícula recorriendo la hélice con posiciones dadas por $r(t)$ es $r'(t) = (-\sin t, \cos t, 1)$.

3.3. Recta tangente a una curva

En esta sección describimos métodos para obtener la recta tangente a una curva en un punto dado.

El procedimiento para curvas planas o espaciales es análogo, difiriendo sólo en el número de ecuaciones implícitas que describen una recta. En el plano $\mathbb{R}^2$ una recta viene dada por una ecuación. En el espacio $\mathbb{R}^3$ existen 3 parámetros de libertad x, y, z luego para obtener un espacio 1-dimensional – e.g. una recta – requerimos de dos condiciones, i.e. dos ecuaciones. Describamos a continuación el caso de curvas planas, el caso de curvas espaciales es análogo.

3.3.1. Curvas planas

Sea C una curva plana parametrizada por $r : \mathbb{R} \rightarrow \mathbb{R}^2$ y $p = r(t_0) \in C$ un punto dado. El vector director de la recta tangente en el punto p es $r'(t_0) \in \mathbb{R}^2$. Dado que tenemos el punto $p = r(t_0)$ y el vector director $r'(t_0)$, podemos escribir la forma paramétrica para la recta tangente $T_p C$:

$$T_p C = \{q \in \mathbb{R}^2 : q = r(t_0) + \lambda r'(t_0), \text{ para algún } \lambda \in \mathbb{R}\}.$$

Las coordenadas de un punto q son (x, y) . Teniendo en cuenta que $r(t_0) = (x(t_0), y(t_0))$ y $r'(t_0) = (x'(t_0), y'(t_0))$ podemos escribir las ecuaciones implícitas en (x, y) para la recta:

$$\begin{aligned} (x, y) = q &= r(t_0) + \lambda r'(t_0) = (x(t_0) + \lambda x'(t_0), y(t_0) + \lambda y'(t_0)) \iff \\ &\iff x = x(t_0) + \lambda x'(t_0), \quad y = y(t_0) + \lambda y'(t_0) \\ &\iff \frac{x - x(t_0)}{x'(t_0)} = \lambda = \frac{y - y(t_0)}{y'(t_0)} \iff \\ &\iff (y'(t_0)/x'(t_0))x + (-x'(t_0)/y'(t_0))y = x(t_0)y'(t_0) - y(t_0)x'(t_0) \end{aligned}$$

Las últimas igualdades son ciertas siempre que $x'(t_0), y'(t_0) \neq 0$.

Existe un método más directo para calcular la recta tangente en el punto $p = r(t_0)$. Recordemos que para hallar la ecuación implícita de una recta en el plano con vector director $v = (a, b)$ basta encontrar un vector w perpendicular a v , hacer el producto escalar $w \cdot (x, y)$ e igualar a una constante adecuada. En el plano podemos tomar $w = (-b, a)$. Deducimos entonces el siguiente:

Método: Recta tangente a una curva plana parametrizada en un punto

Ingredientes: Curva plana $C \subset \mathbb{R}^2$ descrita por $r(t) = (x(t), y(t))$, punto $p = r(t_0) \in C$.

Resultado: La ecuación implícita de la recta $T_p C$ tangente a C en p .

Procedimiento: 1. Calcular $r'(t) = (x'(t), y'(t))$ y sustituir $t = t_0$. Obtenemos $r'(t_0)$.

2. Considerar el vector $w = (-y'(t_0), x'(t_0))$.

3. Calcular el producto escalar: $w \cdot (x, y)$.

4. Calcular la constante $D = w \cdot r(t_0)$.

La ecuación implícita de la recta $T_p C$ es $\{w \cdot (x, y) = D\}$.

Ejemplos: 1. **Segmentos.** Sea $C \subset \mathbb{R}^2$ una recta conteniendo al origen y $p \in C$ un punto cualquiera. Entonces $T_p C$ es precisamente C . Esto se comprueba usando una parametrización $r(t) = (t, mt + n)$ de C . En particular la recta tangente a cualquier punto de un segmento uniendo dos puntos $p, q \in \mathbb{R}^2$ es la única recta conteniendo al punto con vector director $q - p$.

2. **Circunferencias.** Sea C la circunferencia unitada centrada en $(0, 0) \in \mathbb{R}^2$ y consideramos el punto $p = (0, 1) \in C$. Una parametrización es $r(t) = (\cos t, \sin t)$, $0 \leq t \leq 2\pi$, y el punto corresponde a $p = (0, 1) = r(\pi/2)$, i.e. $t_0 = \pi/2$. La dirección de la recta tangente es $r'(t) = (-\sin t, \cos t)$. En el punto $p = r(\pi/2)$ esto es $r'(\pi/2) = (-1, 0)$. El vector w y la constante D son

$$w = (0, -1), \quad D = (0, -1) \cdot r(\pi/2) = (0, -1) \cdot (0, 1) = -1.$$

Luego la ecuación de la recta tangente a C en $p = (0, 1)$ es $\{y = 1\}$.

Sea C una curva expresada en forma implícita: $C = \{(x, y) : f(x, y) = 0\}$. La recta tangente se puede calcular con el siguiente procedimiento:

Método: Recta tangente a una curva plana implícita en un punto

Ingredientes: Curva plana $C \subset \mathbb{R}^2$ descrita por $\{f(x, y) = 0\}$, punto $p = (x_0, y_0) \in C$.

Resultado: La ecuación implícita de la recta $T_p C$ tangente a C en p .

Procedimiento: 1. Calcular $\partial_x f$ y $\partial_y f$. Sustituir (x_0, y_0) : $\partial_x f(x_0, y_0)$ y $\partial_y f(x_0, y_0)$.

2. Calcular la constante $c = \partial_x f(x_0, y_0) \cdot x_0 + \partial_y f(x_0, y_0) \cdot y_0$.

La ecuación implícita de la recta $T_p C$ es $\{\partial_x f(x_0, y_0) \cdot x_0 + \partial_y f(x_0, y_0) \cdot y_0 = c\}$.

Ejemplos: 1. **Rectas.** Sea $l = \{(x, y) : y = 3x + 2\}$ una recta. La geometría indica que la recta tangente a una recta debería ser la misma recta. Hallemos la ecuación de la recta tangente en el punto $(x_0, y_0) = (0, 2)$. La ecuación implícita viene dada por $f(x, y) = 3x + 2 - y$. Tenemos $\partial_x f = 3$ y $\partial_y f = -1$. Sustituir $(0, 2)$ no tiene efecto porque las derivadas parciales son constantes. La constante es $c = 3 \cdot 0 + 2 \cdot (-1) = -2$. La recta tangente tiene ecuación $\{3x - y = -2\}$, es decir, $\{y = 3x + 2\}$. Concluimos que la recta tangente coincide con l . Esto sucede para la recta tangente a una recta en un punto arbitrario.

2. Circunferencias. Sea C la circunferencia unidad centrada en $(0, 0) \in \mathbb{R}^2$ y consideramos el punto $p = (x_0, y_0) = (0, 1) \in C$. Una ecuación implícita es $\{x^2 + y^2 - 1 = 0\}$. Tenemos $\partial_x f = 2x$, $\partial_y f = 2y$. Luego $\partial_x f(0, 1) = 0$ y $\partial_y f(0, 1) = 2$. La constante es $c = 0 \cdot 0 + 2 \cdot 1 = 2$. La ecuación de la recta tangente a C en el punto $p = (0, 1)$ es $\{2y = 2\}$, i.e. $\{y = 1\}$.

3.4. Campos y Fuerzas

Las campos vectoriales permiten una formalización rigurosa de nociones físicas como la fuerza o la velocidad. Una partícula pesada en el espacio crea un campo gravitatorio, e.g. el sol es una estrella de $2 \cdot 10^{30}$ kg que mantiene los planetas del sistema solar orbitando a su alrededor. Dos puntos distintos del sistema solar sienten la atracción gravitatoria en distintas direcciones y magnitudes. Un punto cercano al sol, por ejemplo una hormiga en la superficie de Venus siente una atracción hacia el sol mayor que la atracción que siente la misma hormiga en la superficie terrestre¹. La dirección del campo gravitatorio también es distinta. En los días 21 de Marzo y 23 de Septiembre la Tierra se halla en dos puntos antipodales respecto al Sol: si trazamos la recta desde la posición de la Tierra en un 21 de Marzo a la posición de la Tierra en un 23 de Septiembre, el Sol se encuentra en el punto medio, aproximadamente a $15 \cdot 10^7$ km de la Tierra. La fuerza que siente la Tierra con respecto al Sol en estos dos días es una atracción hacia el Sol, en particular las direcciones de la fuerza en estos dos puntos son opuestas.

La ubicuidad de los campos y fuerzas en disciplinas científicas es visible: el modelo del átomo de Bohr requiere de estos campos, la mecánica cuántica en general, la distribución de las moléculas –e.g. la geometría de los fullerenos– o la mecánica celeste. A continuación sentamos las bases matemáticas para poder operar con estas nociones.

3.4.1. Campos vectoriales en $\mathbb{R}^2$

Sean $(x, y) \in \mathbb{R}^2$ las coordenadas cartesianas del plano. Un campo vectorial en el plano $\mathbb{R}^2$ es la asignación de un vector $F(x, y)$ en cada punto $(x, y) \in \mathbb{R}^2$. Equivalentemente, es una aplicación $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Los vectores en el plano tienen dos componentes, denotemos $F(x, y) = (P(x, y), Q(x, y))$. Es decir, $P(x, y)$ es la primera coordenada del vector $F(x, y)$ y $Q(x, y)$ la segunda. Es habitual encontrar las siguientes notaciones:

$$F(x, y) = P(x, y) \cdot \vec{i} + Q(x, y) \cdot \vec{j}, \quad F(x, y) = P(x, y)dx + Q(x, y)dy.$$

Ejemplos: 1. El campo $F(x, y) = (x, 0)$ asigna un vector horizontal a todo punto (x, y) salvo a lo largo del eje $OY = \{(x, y) : x = 0\}$, donde asigna el vector nulo. El campo $F(x, y) = (0, -y)$ asigna un vector vertical a todo punto (x, y) de módulo $|y|$. El campo constante $F(x, y) = (0, -9,81)$ es una primera aproximación al campo gravitatorio para sistemas físicos cercanos a la superficie terrestre.

2. El campo $F(x, y) = (x, y)$ es el campo radial: en el origen es nulo y es transversal a lo largo de las circunferencias centradas en el origen.

3. El campo $F(x, y) = (-y, x)$ es un campo nulo en el origen y tangente a las circunferencias centradas en el origen.

Podemos restringir campos vectoriales en el plano a lo largo de una curva C . En la subsección 3.2.1 se ha descrito como expresar una curva C como imagen de una función $r : [a, b] \rightarrow \mathbb{R}^2$. El punto $r(t) = (x(t), y(t)) \in \mathbb{R}^2$ es un vector de dos componentes, cada una dependiendo exclusivamente de la variable $t \in [a, b]$. Así pues, un punto del plano $(x, y) \in \mathbb{R}^2$ que esté en la curva C es de la forma $(x, y) = (x(t), y(t))$. Esto es, existe un valor de t de modo que $r(t) = (x, y)$.

Definición 3. La restricción del campo vectorial $F(x, y)$ a la curva $C \subset \mathbb{R}^2$ es la función

$$F|_C : \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto F(r(t)) = F(x(t), y(t)).$$

Ejemplos: 1. Sea C la circunferencia unidad descrita por $r(t) = (\cos t, \sin t)$, $0 \leq t \leq 2\pi$. Tenemos $x(t) = \cos t$, $y(t) = \sin t$. La restricción del campo vectorial

$$F(x, y) = (x^2 + e^y, 3y^4 - xy)$$

¹Las masas de Venus y la Tierra son parecidas y consideramos la atracción de la hormiga hacia cada uno de los dos planetas la misma. Luego la magnitud de la atracción hacia el Sol sólo depende de la distancia a éste.

a la curva C es la función vectorial de una variable

$$F(r(t)) = ((\cos t)^2 + e^{\sin t}, 3(\sin t)^4 - \cos t \sin t).$$

2. Sea $F(x, y) = (y - \ln x, xy^7 + \tan x)$ y C el segmento $r(t) = (t, 3t + 2)$ con $-1 \leq t \leq 1$. Por definición $x(t) = t$, $y(t) = 3t + 2$. La restricción de F al segmento es

$$F(x(t), y(t)) = (3t + 2 - \ln t, t(3t + 2)^7 + \tan t).$$

3.4.2. Campos vectoriales en $\mathbb{R}^3$

Las nociones desarrolladas en esta sección generalizan al espacio la sección 3.4.1. La noción de campo vectorial generaliza la noción de función. Una función definida en el espacio $\mathbb{R}^3$ es una asignación de un valor $f(p)$ a cada punto $p \in \mathbb{R}^3$:

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}, \quad p \longmapsto f(p).$$

Un campo vectorial F asigna un objeto a cada punto $p \in \mathbb{R}^3$, a diferencia de una función este objeto no es un valor real, es un vector $F(p) \in \mathbb{R}^3$:

Definición 4. Un campo vectorial F en el espacio $\mathbb{R}^3$ es una aplicación diferenciable

$$F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3, \quad p \longmapsto F(p) = (F_1(p), F_2(p), F_3(p)).$$

Ejemplos: 1. El campo $F(x, y, z) = (-y, x, z^2)$. En general, dadas tres funciones diferenciables $F_1, F_2, F_3 : \mathbb{R}^3 \longrightarrow \mathbb{R}$ podemos construir el campo vectorial $F = (F_1, F_2, F_3)$.

2. Campo gravitatorios. Un campo gravitatorio es un campo vectorial F definido en un punto $p \in \mathbb{R}^3$ como la fuerza gravitacional ejercida en una masa unidad en el punto p . Sea G la constante gravitatoria, $G \simeq 6,674 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$ y consideramos un objeto de masa m_0 localizado en el origen $(0, 0, 0) \in \mathbb{R}^3$. El campo gravitatorio creado por la masa m_0 en el origen es

$$F(x, y, z) = -Gm_0 \left(\frac{x}{x^2 + y^2 + z^2}, \frac{y}{x^2 + y^2 + z^2}, \frac{z}{x^2 + y^2 + z^2} \right).$$

Escribiendo $r = (x, y, z)$ para el vector posición, este se suele expresar como

$$F = -\frac{Gm_0}{|r|^2} \cdot r.$$

Dos objetos de masas m_0 y m_1 en el espacio crean un campo gravitatorio. Por el principio de superposición, éste es justamente la suma de los campos gravitatorios creados por cada objetos.

3. Un campo $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ es uniforme si el vector asignado $F(p)$ es independiente del punto p . Equivalentemente, las componentes F_1, F_2 y F_3 son funciones constantes. Cerca de la superficie de la tierra, el campo gravitatorio se suele aproximar por el campo uniforme $F(x, y, z) = (0, 0, -9,81)$.

4. Consideremos un objeto con carga eléctrica q en el origen $(0, 0, 0) \in \mathbb{R}^3$. Éste crea un campo eléctrico. Sea ε_0 la permitividad del vacío. Entonces el campo eléctrico creado por el objeto es

$$F(x, y, z) = F(r) = \frac{q}{4\pi\varepsilon_0|r^2|} \cdot r.$$

3.5. Superficies

Una superficie es un objeto bidimensional. Esto es, un conjunto de puntos que es posible recorrer usando movimientos en dos direcciones. Las superficie se suelen ver en el 3-espacio. Si ignoramos el grosor, una hoja de papel, la hoja de un árbol o un mantel son ejemplos de superficies. Existen dos maneras de dar una superficie en el 3-espacio: mediante una parametrización o una ecuación.

3.5.1. Parametrización de una superficie

Una superficie parametrizada es la imagen de un subconjunto del plano vía una aplicación diferenciable al 3-espacio. Siendo precisos:

Una superficie en el espacio $\mathbb{R}^3$ de variables (x, y, z) es la imagen de una aplicación diferenciable

$$r : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \quad t \longmapsto r(u, v) = (x(u, v), y(u, v), z(u, v)), \quad (u, v) \in \Omega.$$

La función r parametriza la superficie.

Ejemplos: 1. **Planos.** Sea $\Omega = \mathbb{R}^2$ y consideramos la parametrización

$$r(u, v) = (u, v, 2), \quad (u, v) \in \mathbb{R}^2.$$

La imagen de la aplicación r es un 2-plano π en $\mathbb{R}^3$. De hecho $\pi = \{z = 2\}$. Si consideramos

$$r(u, v) = (u, v, 3u - 2v + 1), \quad (u, v) \in \mathbb{R}^2,$$

también obtenemos un 2-plano en $\mathbb{R}^3$, distinto de π .

2. **Discos.** Sea $\Omega = \mathbb{D}^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\} \subset \mathbb{R}^2$ un disco en el plano. La imagen de Ω por la parametrización $r(u, v) = (u, v, 0)$ es un disco contenido en el plano $\{z = 0\}$. En este caso la superficie no está curvada. Si en cambio consideramos la parametrización

$$s(u, v) = (u, v, \sqrt{1 - u^2 - v^2}), \quad (u, v) \in \Omega$$

la imagen es la hemisferio norte de la esfera unidad centrada en el origen $0 \in \mathbb{R}^3$. Esta superficie está curvada.

3. **Esferas.** Sea $\Omega = [0, 2\pi] \times [0, \pi] \subset \mathbb{R}^2$. La imagen de Ω por la parametrización

$$r(u, v) = (\cos u \sin v, \sin u \sin v, \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

está contenida en la esfera unidad $\mathbb{S}^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$. Debido al hecho que existen elección de (u, v) de modo que todo punto de $\mathbb{S}^2$ es de la forma $r(u, v)$, la imagen es $\mathbb{S}^2$. Nótese que esta parametrización no es más que la expresión en coordenadas polares con la restricción del radio $\{r = 1\}$.

3.5.2. Ecuación implícita de una superficie

Es posible dar ciertas superficies S mediante una ecuación en el 3-espacio $\mathbb{R}^3$ de variables (x, y, z) . Esto es $S = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = 0\}$. Se dice que la superficie S viene dada en forma implícita.

Ejemplos: 1. **Planos.** Cualquier 2-plano π en $\mathbb{R}^3$ se expresa como $\pi = \{ax + by + cz = d\}$.

2. **Esfera.** La 2-esfera de radio ρ centrada en el origen $0 \in \mathbb{R}^3$ de forma implícita es

$$\mathbb{S}_\rho^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = \rho^2\}.$$

Esta ecuación implícita está en coordenadas cartesianas. Si usamos coordenadas polares $(r, \varphi, \theta) \in (0, \infty) \times [0, 2\pi] \times [0, \pi]$ la ecuación implícita es $\{r = \rho\}$.

3.5.3. Vector Normal a una superficie parametrizada

Sea S una superficie parametrizada por $r : \Omega \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3$. Describamos la construcción de un vector normal a S en un punto cualquiera $p = r(u, v)$, $(u, v) \in \Omega$.

En la sección 3.1 se explica como obtener un vector ortogonal $v \times w$ a dos vectores $v, w \in \mathbb{R}^3$ dados. Apliquemos esto en el punto de nuestra superficie. Un vector normal al punto $p = r(u, v) \in S$ es por definición un vector ortogonal a todos los vectores del plano tangente $T_p S$. El plano $T_p S$ es un espacio lineal dos dimensional y todo vector de éste puede expresarse como suma de múltiplos de dos vectores e_1, e_2 fijados tal que $e_1 \times e_2 \neq 0$. Es decir, si $e_1, e_2 \in T_p S$ con $e_1 \times e_2 \neq 0$ entonces para todo vector

$v \in T_p S$ existen unos únicos valores $\alpha, \beta \in \mathbb{R}$ de modo que $v = \alpha e_1 + \beta e_2$.

Para producir los vectores $e_1, e_2 \in T_p S$ necesitamos dos vectores tangentes a S en el punto p . Supongamos que la superficie S viene descrita por una parametrización regular $r : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$. Por definición, $\partial_u r, \partial_v r \in T_p S$. Así pues, es natural elegir $e_1 = \partial_u r$ y $e_2 = \partial_v r$. En este caso producimos el **vector normal**

$$n(u, v) := \partial_u r \times \partial_v r.$$

En el cálculo del vector normal en un punto particular $p_0 = r(u_0, v_0)$:

$$n(u_0, v_0) = \partial_u r(u_0, v_0) \times \partial_v r(u_0, v_0),$$

donde hemos sustituido los valores (u_0, v_0) antes de hacer el producto vectorial.

Ejemplos: 1. Sea S el trozo de paraboloide descrito por

$$r(u, v) = (u, v, u^2 + v^2) \in S, \quad (u, v) \in [0, 1] \times [0, 1].$$

Calculemos un vector normal en el punto $p = (0, 0, 0) = r(0, 0) \in S$.

Elegimos los vectores tangentes

$$\partial_u r = (1, 0, 2u)|_{(u,v)=(0,0)} = (1, 0, 0), \quad \partial_v r = (0, 1, 2v)|_{(u,v)=(0,0)} = (0, 1, 0).$$

Entonces un vector normal es $n = (1, 0, 0) \times (0, 1, 0) = (0, 0, 1)$.

2. Sea $S \subset \mathbb{R}^3$ la esfera de radio ρ centrada en el origen $(0, 0, 0) \in \mathbb{R}^3$. Calculemos el vector normal en un punto arbitrario $p = (x, y, z) \in S$. Notemos que la geometría de la esfera nos permite deducir que el resultado debe ser un vector en la misma dirección que el vector posición (x, y, z) del punto p . En particular en el punto $p = (\rho, 0, 0) \in S$ el vector normal debe ser proporcional a $n = (1, 0, 0)$.

En primer lugar parametrizamos la esfera usando coordenadas esféricas:

$$r(u, v) = (x, y, z) = (\rho \cos u \sin v, \rho \sin u \sin v, \rho \cos v), \quad (u, v) \in [0, 2\pi] \times (0, \pi).$$

Técnicamente deberíamos escribir $v \in [0, \pi]$, pero la parametrización en coordenadas esféricas no es regular en los polos norte y sur. Esto permite calcular el vector normal en todo punto de la esfera distinto de estos dos. Consideremos los dos vectores tangentes

$$\partial_u r = (-\rho \sin u \sin v, \rho \cos u \sin v, 0), \quad \partial_v r = (\rho \cos u \cos v, \rho \sin u \cos v, -\rho \sin v).$$

El vector normal es $n = \partial_u r \times \partial_v r$,

$$n = (-\rho^2 \cos u \sin^2 v, -\rho^2 \sin u \sin^2 v, -\rho^2 \cos v \sin v) = -\rho^2 \sin v \cdot (x, y, z).$$

En efecto, n es un múltiplo no nulo de (x, y, z) . En particular si $(u, v) = (0, \pi/2)$ obtenemos el punto $p = r(0, \pi/2) = (\rho, 0, 0)$ y el vector normal es $(-\rho^2, 0, 0)$.

3.5.4. Plano tangente a una superficie

La información de la sección anterior nos permite desarrollar un método para el cálculo del plano tangente a una superficie parametrizada en un punto:

Método: Plano tangente a una superficie parametrizada en un punto

Ingredientes: Superficie S descrita por $r(u, v)$, $(u, v) \in \Omega$, punto $p = r(u_0, v_0) \in S$.

Resultado: La ecuación implícita del plano $T_p S$ tangente a S en p .

Procedimiento: 1. Calcular un vector normal en p : $n = \partial_u r(u_0, v_0) \times \partial_v r(u_0, v_0)$.

2. Calcular el producto escalar: $n \cdot (x, y, z)$.

3. Calcular la constante $D = n \cdot r(u_0, v_0)$.

La ecuación implícita del plano $T_p S$ es $\{n \cdot (x, y, z) = D\}$.

Ejemplos: 1. Sea S el trozo de paraboloide descrito por

$$r(u, v) = (u, v, u^2 + v^2) \in S, \quad (u, v) \in [0, 1] \times [0, 1].$$

Calculemos el plano tangente en el punto $p = (0, 0, 0) = r(0, 0) \in S$.

Un vector normal es $n = (1, 0, 0) \times (0, 1, 0) = (0, 0, 1)$. La constante es

$$D = (0, 0, 1) \cdot r(0, 0) = (0, 0, 0) \cdot (0, 0, 1) = 0.$$

Luego la ecuación es $(0, 0, 1) \cdot (x, y, z) = 0$, i.e. $\{z = 0\}$.

2. **Esfera.** Sea $S \subset \mathbb{R}^3$ la esfera de radio ρ centrada en el origen $(0, 0, 0) \in \mathbb{R}^3$. Calculemos el plano tangente en el punto $p = (\rho, 0, 0) \in S$. Usemos la parametrización con coordenadas esféricas, entonces $p = (\rho, 0, 0) = r(0, \pi/2)$. Un vector normal es $n = (-\rho, 0, 0)$. La constante es

$$D = (-\rho, 0, 0) \cdot r(0, \pi/2) = -\rho^2$$

Así pues, el plano tangente $T_p S$ es $n \cdot (x, y, z) = -\rho^2$, i.e. $\{x = \rho\}$.

3. Sea S un plano conteniendo el origen y $p \in S$ arbitrario. El plano tangente $T_p S$ es S .

En caso que la superficie venga dada en forma implícita también podemos usar el siguiente:

Método: Plano tangente a una superficie implícita en un punto

Ingredientes: Superficie $S \subset \mathbb{R}^3$ descrita por $\{f(x, y, z) = 0\}$, punto $p = (x_0, y_0, z_0) \in C$.

Resultado: La ecuación implícita del plano $T_p S$ tangente a S en p .

Procedimiento: 1. Calcular $\partial_x f$, $\partial_y f$ y $\partial_z f$. Sustituir (x_0, y_0, z_0) .

2. Calcular la constante $c = \partial_x f(x_0, y_0, z_0) \cdot x_0 + \partial_y f(x_0, y_0, z_0) \cdot y_0 + \partial_z f(x_0, y_0, z_0) \cdot z_0$.

Ecuación implícita de $T_p S$: $\{\partial_x f(x_0, y_0, z_0) \cdot x_0 + \partial_y f(x_0, y_0, z_0) \cdot y_0 + \partial_z f(x_0, y_0, z_0) \cdot z_0 = c\}$.

1. **Esfera.** Sea $S \subset \mathbb{R}^3$ la esfera de radio ρ centrada en el origen $(0, 0, 0) \in \mathbb{R}^3$. Una ecuación implícita es

$$\mathbb{S}_\rho^2 = \{(x, y, z) : x^2 + y^2 + z^2 = \rho^2\}.$$

Calculemos el plano tangente en el punto $p = (\rho, 0, 0) \in S$. Tenemos $f(x, y, z) = x^2 + y^2 + z^2 - \rho^2$, luego:

$$\partial_x f = 2x, \quad \partial_y f = 2y, \quad \partial_z f = 2z.$$

En el punto $p = (\rho, 0, 0)$ obtenemos

$$\partial_x f(\rho, 0, 0) = 2\rho, \quad \partial_y f(\rho, 0, 0) = 0, \quad \partial_z f(\rho, 0, 0) = 0.$$

Calculamos la constante $c = 2\rho \cdot \rho$. Así pues la ecuación del plano tangente $T_p \mathbb{S}_\rho^2$ es

$$T_p S = \{2\rho x = 2\rho^2\} = \{x = \rho\}.$$

4. Integración

Las nociones básicas de geometría en el plano y el espacio han sido introducidas en la primera parte. En esta segunda parte definiremos operaciones sobre estos objetos. Estas operaciones nos permitirán calcular magnitudes como la longitud de una curva o el área de una superficie. A su vez, la integración da un fundamento riguroso a la noción del trabajo de una fuerza a lo largo de una curva o el flujo de un campo a través de una superficie.

4.1. Integrales en curvas

En esta sección se define la longitud y la integral de una función vectorial a lo largo de una curva. La denominación *integral de línea* proviene del hecho que las curvas son objetos uno dimensionales.

4.1.1. Longitud de una curva plana

Sea $C \subset \mathbb{R}^2$ una curva parametrizada por $r : [a, b] \rightarrow \mathbb{R}^2$. En física la variable $t \in [a, b]$ suele interpretarse como el tiempo y $r(t_0) \in \mathbb{R}^2$ es la posición de una partícula en el plano a tiempo t_0 . Teniendo en cuenta que la velocidad $v(t_0)$ de la partícula a tiempo t_0 es la variación de la posición en el instante t_0 , tenemos $v(t_0) = r'(t_0)$. La longitud del camino recorrido por una partícula en el intervalo temporal $[a, b]$ es la suma de las longitudes infinitesimales. La variación de la longitud equivale al producto de la variación de velocidad por variación del tiempo. Luego para obtener la longitud L debemos sumar $|r'(t)| \cdot dt$ a lo largo de $t \in [a, b]$, esto es

$$L = \int_a^b |r'(t)| dt.$$

Definición 5. Sea C una curva plana descrita por $r : [a, b] \rightarrow \mathbb{R}^2$. La longitud $L(C)$ de la curva plana C es

$$L(C) = \int_a^b |r'(t)| dt.$$

Es sencillo comprobar que para otra parametrización $s : [c, d]$ con la misma curva imagen C se tiene

$$\int_c^d |s'(t)| dt = \int_a^b |r'(t)| dt.$$

Luego la cantidad $L(C)$ es independiente de la parametrización y sólo depende de C .

Ejemplo: 1. Longitud de Segmentos. Sean $p, q \in \mathbb{R}^2$ dos puntos en el plano y $r(t) = t \cdot q + (1 - t) \cdot p$, $0 \leq t \leq 1$ una parametrización del segmento C que les une. La longitud debería ser el módulo del vector $q - p$. En efecto,

$$L(C) = \int_0^1 |r'(t)| dt = \int_0^1 |q - p| dt = |q - p| \int_0^1 dt = |q - p|.$$

2. Longitud de la circunferencia de radio ρ . Sea $C \subset \mathbb{R}^2$ la circunferencia de radio ρ centrada en el origen $(0, 0)$. Una posible parametrización es

$$r(t) = (\rho \cos t, \rho \sin t), \quad 0 \leq t \leq 2\pi.$$

La longitud de esta circunferencia es entonces

$$L(C) = \int_0^{2\pi} |r'(t)| dt = \int_0^{2\pi} |(-\rho \sin t, \rho \cos t)| dt = \int_0^{2\pi} \rho dt = 2\pi\rho.$$

3. Longitud de Elipses. Sea $C \subset \mathbb{R}^2$ la elipse de semiejes $a, b > 0$ centrada en el origen con la parametrización

$$r(t) = (a \cos t, b \sin t), \quad 0 \leq t \leq 2\pi.$$

Luego la longitud de esta elipse es

$$\begin{aligned} L(C) &= \int_0^{2\pi} |r'(t)| dt = \int_0^{2\pi} |(-a \sin t, b \cos t)| dt = \int_0^{2\pi} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = \\ &= 4 \int_0^{\pi/2} |(-a \sin t, b \cos t)| dt = 4 \int_0^{\pi/2} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = 4a \int_0^{\pi/2} \sqrt{1 - \varepsilon \cos^2 t} dt \end{aligned}$$

donde $\varepsilon = \sqrt{1 - b^2/a^2}$ es la excentricidad de la elipse. La integral de la función $\sqrt{1 - \varepsilon \cos^2 t}$ no puede expresarse en términos de funciones elementales. La longitud de una elipse es una cantidad difícil de calcular.

4.1.2. Integral de línea en el plano

La noción geométrica de la integral es la suma infinitesimal de valores. La integral de línea de un campo vectorial a lo largo de una curva es la suma infinitesimal de un valor en cada punto de la curva. Sea $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un campo vectorial y C una curva descrita por $r : [a, b] \rightarrow \mathbb{R}^2$.

En un punto de la curva $r(t) = (x(t), y(t)) \in C$, el campo vectorial F asigna el vector $F(r(t)) = F(x(t), y(t))$. Dado que debemos obtener un valor, i.e. un número, uno podría pensar que debemos considerar el módulo $\|F(r(t))\|$. No es así. El valor que debemos considerar es **el módulo del campo F proyectado a la curva C** . Describamos la proyección a una curva:

Definición 6. Sea C una curva descrita por $r : [a, b] \rightarrow \mathbb{R}^2$, $r(t) = (x(t), y(t))$. Sea $v \in \mathbb{R}^2$ un vector apoyado en el punto $r(t_0) \in \mathbb{R}^2$, la proyección de v a la curva C es el producto escalar $v \cdot r'(t_0)$.

En particular, la proyección del campo vectorial $F(x, y)$ a lo largo de C es $F(r(t)) \cdot r'(t)$.

Método: Cálculo de la proyección de un campo vectorial a una curva

Ingredientes: Campo vectorial $F(x, y)$ en el plano, curva C descrita por $r(t) = (x(t), y(t))$.

Resultado: Una función $f(t)$ de una variable t calculando la magnitud del campo F en la dirección de la curva C .

Procedimiento: 1. Sustituir x e y en $F(x, y)$ por $x(t)$ e $y(t)$. Obtenemos $F(r(t))$.

2. Calculamos la velocidad de la curva: $r'(t) = (x'(t), y'(t))$.

3. Calcular el producto escalar: $F(r(t)) \cdot r'(t)$.

El resultado es este producto escalar $f(t) := F(r(t)) \cdot r'(t)$.

Ejemplo: Calculemos la restricción del campo $F(x, y) = (y^2 + y, 2xy - e^{2y})$ a lo largo de la circunferencia C descrita por $r(t) = (\cos t, \sin t)$. Es decir, $x(t) = \cos t$, $y(t) = \sin t$. Según el procedimiento descrito tenemos:

1. Sustituimos x e y por $x(t) = \cos t$, $y(t) = \sin t$ en $F(x, y)$:

$$F(r(t)) = ((\sin t)^2 + \sin t, 2 \cos t \sin t - e^{2 \sin t}).$$

2. La velocidad de la curva es $r'(t) = (x'(t), y'(t)) = (-\sin t, \cos t)$.

3. El producto escalar del campo con la velocidad es

$$\begin{aligned} F(r(t)) \cdot r'(t) &= ((\sin t)^2 + \sin t, 2 \cos t \sin t - e^{2 \sin t}) \cdot (-\sin t, \cos t) = \\ &= -\sin t \cdot ((\sin t)^2 + \sin t) + \cos t \cdot (2 \cos t \sin t - e^{2 \sin t}). \end{aligned}$$

La integral de línea del campo F a lo largo de C es la **integral de la función de una variable** $F(r(t)) \cdot r'(t)$. Siendo precisos:

Definición 7. Sea $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un campo vectorial y $C \subset \mathbb{R}^2$ una curva descrita por $r(t)$ con $a \leq t \leq b$. La integral de línea de F a lo largo de C es

$$\int_C F \cdot dl := \int_a^b F(r(t)) \cdot r'(t) dt$$

Ejemplo: 1. Calculemos la integral de línea del campo vectorial $F(x, y) = (y, x^2)$ en la curva C dada por $r(t) = (4 - t, 4t - t^2)$ con $0 \leq t \leq 3$. Según la descripción de la curva C : $x(t) = 4 - t$, $y(t) = 4t - t^2$. Luego

$$F(r(t)) = (4t - t^2, (4 - t)^2), \quad r'(t) = (-1, 4 - 2t), \quad F(r(t)) \cdot r'(t) = t^2 - 4t + (4 - 2t)(4 - t)^2$$

La integral de línea de F a lo largo de C es

$$\int_C F \cdot dl = \int_0^3 t^2 - 4t + (4 - 2t)(4 - t)^2 dt = \frac{69}{2}$$

2. Calculemos la integral de línea del campo vectorial $F(x, y) = (x^3 - 2x^2, x - y/2)$ al movernos en la curva C descrita por $r(t) = (t, t^2)$ con $-1 \leq t \leq 2$. En este caso $r(t) = (x(t), y(t)) = (t, t^2)$ y tenemos

$$F(r(t)) = (t^3 - 2t^2, t - t^2/2), \quad r'(t) = (1, 2t) \implies F(r(t)) \cdot r'(t) = t^3 - 2t^2 + 2t^2 - t^3 = 0$$

Luego la integral de línea de F a lo largo de C es

$$\int_C F \cdot dl = \int_{-1}^2 0 dt = 0$$

En general si el campo $F(r(t))$ es perpendicular al vector tangente $r'(t)$, i.e. $F(r(t)) \cdot r'(t) = 0$, la integral es cero.

4.1.3. Teorema de Green

En esta sección establecemos el Teorema de Green. Este resultado aplica sólo a integrales de línea a lo largo de una **curva plana cerrada**. Sea F un campo vectorial en el plano $\mathbb{R}^2$. Tal y como describimos en la sección 3.4.1 expresamos este campo como

$$F(x, y) = (P(x, y), Q(x, y)) = P(x, y)dx + Q(x, y)dy$$

Sea $C \subset \mathbb{R}^2$ una curva plana cerrada y orientada. Esta curva es 1-dimensional y la integral de línea de F a lo largo de C es equivalente al cálculo de una integral en 1 variable. Consideremos el dominio Ω definido como la componente conexa acotada de las dos componentes en las que C separa el plano $\mathbb{R}^2$. La frontera de Ω es C , i.e. $\partial\Omega = C$.

El resultado siguiente permite calcular la integral de línea a lo largo de C como una integral en dos variables:

Teorema 4.1. (Teorema de Green) Sea Ω una dominio en el plano con frontera una curva conexa orientada cerrada C . Dado un campo vectorial $F(x, y) = (P(x, y), Q(x, y))$,

$$\int_C F \cdot dl = \iint_{\Omega} (\partial_x Q - \partial_y P) dx dy$$

Observación: Para aplicar el teorema **no** es necesario tener la curva C descrita por $r(t)$. Sólo es necesario describir el dominio interior Ω .

Ejemplos: 1. **Circunferencias.** Sea C la circunferencia de radio ρ centrada en el origen. El dominio Ω es el disco de radio ρ centrado en el origen: $\Omega = \{(r, \theta) : r = \rho\} = \{(x, y) : x^2 + y^2 = \rho^2\}$. Consideramos $F(x, y) = (P(x, y), Q(x, y)) = (y^2 + y + e^x, 2xy - e^{2y} + \tan y)$. La integral de línea de F a lo largo de C es

$$\int_C F \cdot dl = \iint_{\Omega} Q_x - P_y dx dy = \iint_{\Omega} -1 \cdot dx dy = -\text{Area}(\Omega) = -\pi\rho^2.$$

2. Sea $C \subset \mathbb{R}^2$ el triángulo de vértices $(-1, 0)$, $(1, 0)$, $(0, 1)$ y $F(x, y) = (e^y, -\sin(\pi x))$. El dominio Ω es el interior del triángulo, cuyos límites de integración se pueden tomar como $0 \leq y \leq 1$ y $y - 1 \leq x \leq 1 - y$. Vía el teorema de Green, la integral de línea de F a lo largo de C se puede calcular integrando $Q_x - P_y$ en el interior del triángulo:

$$I = \int_{\Omega} (Q_x - P_y) \cdot dx dy = \int_0^1 \int_{y-1}^{1-y} (-\pi \cos(\pi x) - e^y) \cdot dx dy = -2(e - 2) - \frac{4}{\pi}.$$

3. Sea $C \subset \mathbb{R}^2$ una curva plana cerrada cualquiera y $F = (P, Q)$ una fuerza con $Q_x = P_y$. Entonces

$$\int_C F \cdot dl = 0.$$

Instancias de este hecho son las fuerzas

$$F(x, y) = (x, 0), F(x, y) = (0, y), F(x, y) = (x, y) \text{ y } F(x, y) = (y, x).$$

Ver Ejercicio 7 de T2.2.6.

4.1.4. Longitud de una curva en el espacio

El concepto de longitud para una curva en el espacio coincide con el de longitud de una curva plana. La única variación es que la imagen de $r(t)$ está contenida en $\mathbb{R}^3$:

Definición 8. Sea $C \subset \mathbb{R}^3$ una curva espacial descrita por $r : [a, b] \rightarrow \mathbb{R}^3$. La longitud $L(C)$ de la curva plana C es

$$L(C) = \int_a^b |r'(t)| dt.$$

Dada otra parametrización $s : [c, d]$ con la misma curva imagen C se tiene

$$\int_c^d |s'(t)| dt = \int_a^b |r'(t)| dt,$$

y la longitud $L(C)$ es independiente de la parametrización.

Ejemplo: 1. Longitud de Segmentos. Sean $p, q \in \mathbb{R}^3$ dos puntos en el espacio y $r(t) = t \cdot q + (1-t) \cdot p$, $0 \leq t \leq 1$ una parametrización del segmento C que les une. La longitud es el módulo del vector $q - p$:

$$L(C) = \int_0^1 |r'(t)| dt = \int_0^1 |q - p| dt = |q - p| \int_0^1 dt = |q - p|.$$

2. Longitud de una hélice. Sea $C \subset \mathbb{R}^3$ la hélice descrita por

$$r : [0, 2\pi] \rightarrow \mathbb{R}^3, \quad r(t) = (\cos t, \sin t, t).$$

La velocidad es $r'(t) = (-\sin t, \cos t, 1)$. Luego la longitud es

$$L(C) = \int_0^{2\pi} \|(-\sin t, \cos t, 1)\| dt = \int_0^{2\pi} \left(\sqrt{\sin^2 t + \cos^2 t + 1^2} \right) dt = \int_0^{2\pi} \sqrt{2} dt = 2\sqrt{2}\pi.$$

4.1.5. Integral de línea en el espacio

Esta noción es el análogo 3-dimensional de 4.1.2. La integral de línea de un campo vectorial F en una curva es la suma infinitesimal de un valor asociada a F en cada punto de la curva. Sea $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vectorial y C una curva descrita por $r : [a, b] \rightarrow \mathbb{R}^3$.

En un punto de la curva $r(t) = (x(t), y(t), z(t)) \in C$, el campo vectorial F asigna el vector $F(r(t)) = F(x(t), y(t), z(t))$. El valor que se integra a lo largo de la curva es **el módulo del campo F proyectado a la curva C** . Recordemos que la proyección del campo vectorial $F(x, y, z)$ a lo largo de C es $F(r(t)) \cdot r'(t)$.

La integral de línea del campo F en C es la integral de la función de una variable $F(r(t)) \cdot r'(t)$ con $a \leq t \leq b$.

Definición 9. Sea $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un campo vectorial y $C \subset \mathbb{R}^3$ una curva descrita por $r(t)$ con $a \leq t \leq b$. La integral de línea de F a lo largo de C es

$$\int_C F \cdot dl := \int_a^b F(r(t)) \cdot r'(t) dt$$

Nótese que la definición es idéntica al caso 2-dimensional establecido en la sección 4.1.2.

4.2. Integrales de Superficie

Esta sección introducimos la integración de una función escalar a lo largo de una superficie. En el caso particular que la función escalar es constante igual a 1 obtendremos el área de una superficie.

4.2.1. Área de una superficie

La magnitud $\|v \times w\|$ del producto vectorial de dos vectores $v, w \in \mathbb{R}^3$ es el área del paralelogramo generado por v, w en el plano $\langle v, w \rangle$. Sea S una superficie parametrizada por $r : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$. El área de la superficie S se puede entender como la suma infinitesimal de las áreas infinitesimales de los paralelogramos generados por $\partial_u r$ y $\partial_v r$ en el plano tangente a cada punto $p \in S$.

Definición 10. Sea S una superficie descrita por $r : \Omega \rightarrow \mathbb{R}^3$. El área $A(S)$ de la superficie S es

$$A(S) = \iint_{\Omega} |\partial_u r \times \partial_v r| du dv.$$

Ejemplos: 1. **Esferas.** Sea $\Omega = [0, 2\pi] \times [0, \pi] \subset \mathbb{R}^2$. La imagen de Ω por la parametrización

$$r(u, v) = (\rho \cos u \sin v, \rho \sin u \sin v, \rho \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

es la esfera $\mathbb{S}_{\rho}^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ de radio ρ centrada en el origen. Tal y como se ha calculado anteriormente el vector normal es

$$n = (-\rho^2 \cos u \sin^2 v, -\rho^2 \sin u \sin^2 v, -\rho^2 \cos v \sin v) = -\rho^2 \sin v \cdot (x, y, z).$$

El área de $\mathbb{S}_{\rho}^2$ es la integral del módulo del vector normal a lo largo de S :

$$\begin{aligned} A(S) &= \iint_{\Omega} \|n\| du dv = \int_0^{2\pi} \int_0^{\pi} |-\rho^2 \sin v \cdot| du dv = \\ &= \rho^2 \int_0^{2\pi} \int_0^{\pi} |\sin v| \cdot du dv = 2\pi \rho^2 \int_0^{\pi} |\sin v| dv = 4\pi \rho^2. \end{aligned}$$

4.2.2. Integral de superficie

Sea S una superficie parametrizada por $r : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ y $f : S \rightarrow \mathbb{R}$ una función escalar. La integral de la función f a lo largo de S es la suma de las área ponderadas con la densidad dada por f :

Definición 11. Sea S una superficie descrita por $r : \Omega \rightarrow \mathbb{R}^3$. La integral de f a lo largo de la superficie S es

$$\iint_{\Omega} f(u, v) |\partial_u r \times \partial_v r| du dv.$$

Ejemplos: 1. **Esferas.** Consideramos $S = \mathbb{S}_{\rho}^2$ con la parametrización

$$r(u, v) = (\rho \cos u \sin v, \rho \sin u \sin v, \rho \cos v), \quad (u, v) \in [0, 2\pi] \times [0, \pi].$$

y la función $f : \mathbb{S}_{\rho}^2 \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $f(x, y, z) = z$. Restringida a la superficie $\mathbb{S}_{\rho}^2$ esta función es

$$f|_S = (f \circ r)(u, v) = \rho \cos v.$$

Teniendo en cuenta que la función f tiene valores opuestos en el hemisferio norte y el hemisferio sur, la integral debería ser cero. El vector normal es

$$n = (-\rho^2 \cos u \sin^2 v, -\rho^2 \sin u \sin^2 v, -\rho^2 \cos v \sin v) = -\rho^2 \sin v \cdot (x, y, z).$$

La integral de f a lo largo de S es

$$\begin{aligned} \iint_{\Omega} f \|n\| du dv &= \int_0^{2\pi} \int_0^{\pi} \rho \cos v |-\rho^2 \sin v \cdot| du dv = \\ &= \rho^2 \int_0^{2\pi} \int_0^{\pi} \rho \cos v |\sin v| \cdot du dv = 2\pi \rho^3 \int_0^{\pi} \cos v \sin v dv = 0. \end{aligned}$$

Algunos Operadores Diferenciales Importantes

Roberto Giménez Conejero

17 de marzo de 2025

1. Operadores básicos

Estas notas incluyen una presenciación de los operadores básicos que usaremos durante este curso. Diferenciaremos entre los que actúan sobre campos escalares de los que actúan sobre campos vectoriales. Son muy importantes en el estudio de fenómenos electromagnéticos; si quieres saber más sigue este enlace [AQUI](#).

1. OPERADORES QUE ACTÚAN SOBRE CAMPOS ESCALARES

Consideremos un sistema de coordenadas $(x_1, \dots, x_n)$ en $\mathbb{R}^n$. Si Ω es un recinto en $\mathbb{R}^n$, y $f : \Omega \rightarrow \mathbb{R}$ es un campo escalar, entonces definimos los siguientes operadores diferenciales.

■ El operador gradiente

El gradiente de f es $\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$. En consecuencia, el operador gradiente se define como:

$$\nabla := \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n} \right). \quad (1)$$

Puedes saber más sobre este operador: [AQUI](#)

■ El operador laplaciano

El laplaciano de f es $\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$. En consecuencia, el operador laplaciano se define como

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}, \quad \text{o también } \Delta = \nabla^2. \quad (2)$$

Observa que este operador es de segundo orden, es decir intervienen las derivadas segundas para definirlo. Puedes saber más sobre este operador: [AQUI](#)

2. OPERADORES QUE ACTÚAN SOBRE CAMPOS VECTORIALES

Consideremos un sistema de coordenadas (x_1, x_2, x_3) en $\mathbb{R}^3$. Si Ω es un recinto en $\mathbb{R}^3$, y $\vec{F} : \Omega \rightarrow \mathbb{R}^3$, $\vec{F} = (F_1, F_2, F_3)$, es un campo escalar en Ω , entonces definimos los siguientes operadores diferenciales.

■ El operador divergencia

La divergencia de $\vec{F}$ es el campo escalar $\text{div}(\vec{F}) = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}$. En consecuencia, el operador divergencia se define como:

$$\text{div} := \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3}, \quad \text{o también } \text{div}(\vec{F}) = \nabla \cdot \vec{F} \quad (3)$$

Puedes saber más sobre este operador: [AQUI](#)

■ El operador rotacional

El rotacional de $\vec{F}$ es el campo vectorial

$$\text{rot}(\vec{F}) = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \right) \vec{i} + \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1} \right) \vec{j} + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \vec{k}.$$

En consecuencia, el operador rotacional se define como:

$$\text{rot}(\vec{F}) = \nabla \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ F_1 & F_2 & F_3 \end{vmatrix}. \quad (4)$$

Puedes saber más sobre este operador: [AQUI](#)

2. El álgebra de los operadores diferenciales

Algunas identidades notables entre operadores son las siguientes.

■ Aplicación sobre productos.

$$\begin{aligned} \nabla(\phi\psi) &= \nabla(\phi)\psi + \nabla(\psi)\phi, & \nabla \cdot (\phi\vec{F}) &= \nabla(\phi) \cdot \vec{F} + \phi \nabla \cdot \vec{F}, \\ \nabla \times (\vec{F} \times \vec{G}) &= (\nabla \times \vec{F}) \cdot \vec{G} - (\nabla \times \vec{G}) \cdot \vec{F}, & \nabla \times (\phi\vec{F}) &= \nabla(\phi) \times \vec{F} + \phi \nabla \times \vec{F}. \end{aligned}$$

■ Identidades de Green:

1. $\nabla \cdot (\phi \nabla \psi) = \nabla(\phi) \cdot \nabla(\psi) + \phi \nabla^2 \psi.$

2. $\nabla \cdot (\phi \nabla \psi - \psi \nabla \phi) = \phi \nabla^2 \psi - \psi \nabla^2 \phi.$

Teoremas Fundamentales del Cálculo Vectorial

Maria-Angeles Zurro

17 de marzo de 2025

1. Introducción

Estas notas incluyen una presenciación de los teoremas básicos del cálculo vectorial. Son tres; el teorema de Green (Teorema 2.1) para recintos de $\mathbb{R}^2$, y los teoremas de Stokes (Teorema 3.1) y de la divergencia (Teorema 4.1) para recintos en $\mathbb{R}^3$. Los tres relacionan la integración sobre un recinto con la integración en la frontera del mismo. También es este el caso del Teorema Fundamental del Cálculo para funciones de una variable real, ya que si $I = [a, b]$ entonces su frontera es $\partial I = \{a, b\}$ y el resultado es

$$\int_I f(t)dt = \sum_{x \in \{b, a\}} F(x) = F(b) - F(a),$$

donde F es una primitiva de f . Observa que en la fórmula anterior hay implícito un orden en ∂I ya que $a < b$. Este concepto se generaliza a más dimensiones en el concepto de *orientación del borde* del recinto considerado, y será estudiado en cada teorema.

2. Teorema de Green. Formulación general

Teorema 2.1. (*Teorema de Green*) Sea $\mathcal{D}$ un dominio de $\mathbb{R}^2$ cuya frontera $\partial\mathcal{D}$ es una curva cerrada simple, horientada en el sentido antihorario. Sea $\vec{F} = (F_1, F_2)$ un campo sobre un abierto que contiene a $\mathcal{D} \cup \partial\mathcal{D}$. Entonces

$$\oint_{\partial\mathcal{D}} F_1 dx + F_2 dy = \iint_{\mathcal{D}} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA . \quad (1)$$

Ejemplo 2.1. Sea $\mathcal{D}$ el triángulo de $\mathbb{R}^2$ de vértices $(0, 0)$, $(2, 0)$ y $(2, 2)$. Calcula la circulación del campo $\vec{F} = (\sin(x), x^2y^3)$ a lo largo de la trayectoria circular $\mathcal{C} = \partial\mathcal{D}$.

Para calcular esta circulación basta usar el Teorema de Green y por esto calcular

$$\iint_{\mathcal{D}} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA.$$

En concreto, como $F_1 = \sin(x)$ y $F_2 = x^2y^3$, y el recinto se describe como

$$\mathcal{D} = \{ (x, y) | 0 \leq x \leq 2, 0 \leq y \leq x \},$$

la circulación pedida se calcula como

$$\int_0^2 \left(\int_0^x (2xy^3 - 0) dy \right) dx = \frac{16}{3}.$$

Observa que si parametrizamos el borde del triángulo con la orientación pedida la complejidad de la integral aumenta. Una posible parametrización sería: para $t \in [0, 3]$ definimos

$$\vec{c}(t) = \begin{cases} (1-t)(0, 0) + t(2, 0) & 0 \leq t \leq 1, \\ (2-t)(2, 0) + (t-1)(2, 2) & 1 \leq t \leq 2, \\ (3-t)(2, 2) + (t-2)(0, 0) & 2 \leq t \leq 3. \end{cases}$$

Nota 2.1. El Teorema de Green es válido en dominios más generales. La frontera en general puede estar formada por más de una curva cerrada simple, pongamos $\partial\mathcal{D} = \mathcal{C}_1 \cup \dots \cup \mathcal{C}_s$, y su enunciado general es

$$\oint_{\partial\mathcal{D}} \vec{F} \cdot d\vec{\ell} = \iint_{\mathcal{D}} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dA.$$

3. Teorema de Stokes

Para un campo $\vec{F} = (F_1, F_2, F_3)$ definimos el campo rotacional de $\vec{F}$ y lo denotamos por $\text{rot}(\vec{F})$ como el campo

$$\begin{aligned} \text{rot}(\vec{F}) = \nabla \times \vec{F} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial_x & \partial_y & \partial_z \\ F_1 & F_2 & F_3 \end{vmatrix} = \\ &= (\partial_y F_3 - \partial_z F_2) \vec{i} + (\partial_z F_1 - \partial_x F_3) \vec{j} + (\partial_x F_2 - \partial_y F_1) \vec{k}. \end{aligned} \quad (2)$$

El teorema de Stokes es la extensión a dimensión 3 del Teorema de Green [2.1](#). Se formula de la siguiente manera.

Teorema 3.1. (Teorema de Stokes) Sea $\mathcal{S}$ una superficie de $\mathbb{R}^3$ cuya frontera $\partial\mathcal{S}$ es una unión finita de curvas cerradas simples, orientada positivamente respecto de la normal exterior a $\mathcal{S}$. Sea $\vec{F} = (F_1, F_2, F_3)$ un campo sobre un abierto que contiene a $\mathcal{S} \cup \partial\mathcal{S}$. Entonces

$$\oint_{\partial\mathcal{S}} \vec{F} \cdot d\vec{\ell} = \iint_{\mathcal{S}} \text{rot}(\vec{F}) \cdot d\vec{S} . \quad (3)$$

Ejemplo 3.1. Calcula la circulación del campo $\vec{F} = (\sin(x), e^y + x^2, z^4 + 2x^2)$ sobre la frontera del triángulo de vértices $A = (3, 0, 0)$, $B = (0, 2, 0)$ y $C = (0, 0, 1)$.

Para calcular esta circulación basta usar el Teorema de Stokes y por esto calcular

$$\iint_{\mathcal{S}} \text{rot}(\vec{F}) \cdot d\vec{S}$$

En concreto, como $F_1 = \sin(x)$, $F_2 = x^2 + e^y$, $F_3 = 2x^2 + z^4$, su rotacional es

$$\text{rot}(\vec{F}) = \nabla \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \partial_x & \partial_y & \partial_z \\ \sin(x) & x^2 + e^y & 2x^2 + z^4 \end{vmatrix} = (-4x)\vec{j} + (2x)\vec{k}.$$

Para calcular la integral de superficie tenemos que parametrizar la superficie $\mathcal{S}$. En este caso, como el triángulo está contenido en el plano $-2x - 3y + 3z + 6 = 0$ tenemos que la aplicación de parametrización G está definida sobre el triángulo plano Ω de vértices $P = (0, 0)$, $Q = (3, 0)$ y $R = (0, 2)$ por la fórmula:

$$G(x, y) = \left(x, y, \frac{-6 + 2x + 3y}{3} \right), \quad (4)$$

para (x, y) en el triángulo

$$\Omega = \{(x, y) | 0 \leq x \leq 3, 0 \leq y \leq -\frac{2}{3}x + 2\}.$$

El vector normal asociado a esta parametrización se calcula como sigue. En primer lugar calculamos las derivadas parciales de la parametrización:

$$\partial_x G = (1, 0, \frac{2}{3}), \quad \partial_y G = (0, 1, 1), \quad (5)$$

lo que da el campo normal a la superficie

$$\vec{n}(x, y) = \partial_x G \times \partial_y G = \left(-\frac{2}{3}, -1, 1 \right). \quad (6)$$

Observa que está orientado positivamente porque su componente z , $z = 1$, es siempre positiva. Con esta orientación se tiene que

$$\iint_{\mathcal{S}} \text{rot}(\vec{F}) \cdot d\vec{S} = \int_{\Omega} (0, -4x, 2x) \cdot \left(-\frac{2}{3}, -1, 1\right) dx dy = \int_{\Omega} (6x) dx dy.$$

Esta integral se calcula usando el teoma de Fubini de integración iterada para dar:

$$\int_{\Omega} (6x) dx dy = 6 \int_0^3 \int_0^{-\frac{2}{3}x+2} x dy dx = 6 \int_0^3 x \left(-\frac{2}{3}x + 2\right) dx = 18.$$

En consecuencia, el valor de la circulación pedida es de $18u$.

4. Teorema de la divergencia

Para un campo $\vec{F} = (F_1, F_2, F_3)$ definimos el campo escalar divergencia de $\vec{F}$ y lo denotamos por $\text{div}(\vec{F})$ como el campo

$$\text{div}(\vec{F}) = \nabla \cdot \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}. \quad (7)$$

Nota 4.1. La tasa de circulación a través de $\mathcal{S}$ se calcula usando este operador div mediante el teorema de la divergencia.

El teorema de de la divergencia es la extensión a dimensión 3 ("sólidos") del Teorema de Stokes 3.1. Se formula de la siguiente manera.

Teorema 4.1. (Teorema de la divergencia) Sea $\mathcal{S}$ una superficie cerrada que limita una región $\mathcal{W}$ en $\mathbb{R}^3$. Supongamos que $\mathcal{S}$ es suave a trozos y orientada por vectores normales que apuntan al exterior de $\mathcal{W}$. Sea $\vec{F}$ un campo vectorial cuyo dominio contiene a $\mathcal{W}$. Entonces

$$\iint_{\mathcal{S}} \vec{F} \cdot d\vec{S} = \iiint_{\mathcal{W}} \text{div}(\vec{F}) dV. \quad (8)$$

Ejemplo 4.1. Calcula la integral $\iint_{\mathcal{S}} \vec{F} \cdot d\vec{S}$ donde $\mathcal{S}$ es la frontera de la caja $\mathcal{W} = [0, 2] \times [0, 3] \times [0, 1]$, y el campo viene dado por $\vec{F} = (x^2, z^4, e^z)$.

En primer lugar calcularemos la divergencia del campo dado. En concreto

$$\text{div}(\vec{F}) = \nabla \cdot \vec{F} = \frac{\partial x^2}{\partial x} + \frac{\partial z^4}{\partial y} + \frac{\partial e^z}{\partial z} = 2x + e^z.$$

Por otro lado, el sólido $\mathcal{W}$ ya tiene una descripción adecuada para usar el teorema de Fubini. En consecuencia,

$$\begin{aligned} \iiint_{\mathcal{W}} \operatorname{div}(F) \, dV &= \int_0^2 \left(\int_0^3 \left(\int_0^1 (2x + e^z) \, dz \right) dy \right) dx = \\ &\int_0^2 \left(\int_0^3 (2x + e - 1) \, dy \right) dx = 6e + 6. \end{aligned}$$

Luego el flujo pedido es de $(6e + 6)u.$ en la dirección de la normal exterior a la frontera de la caja.