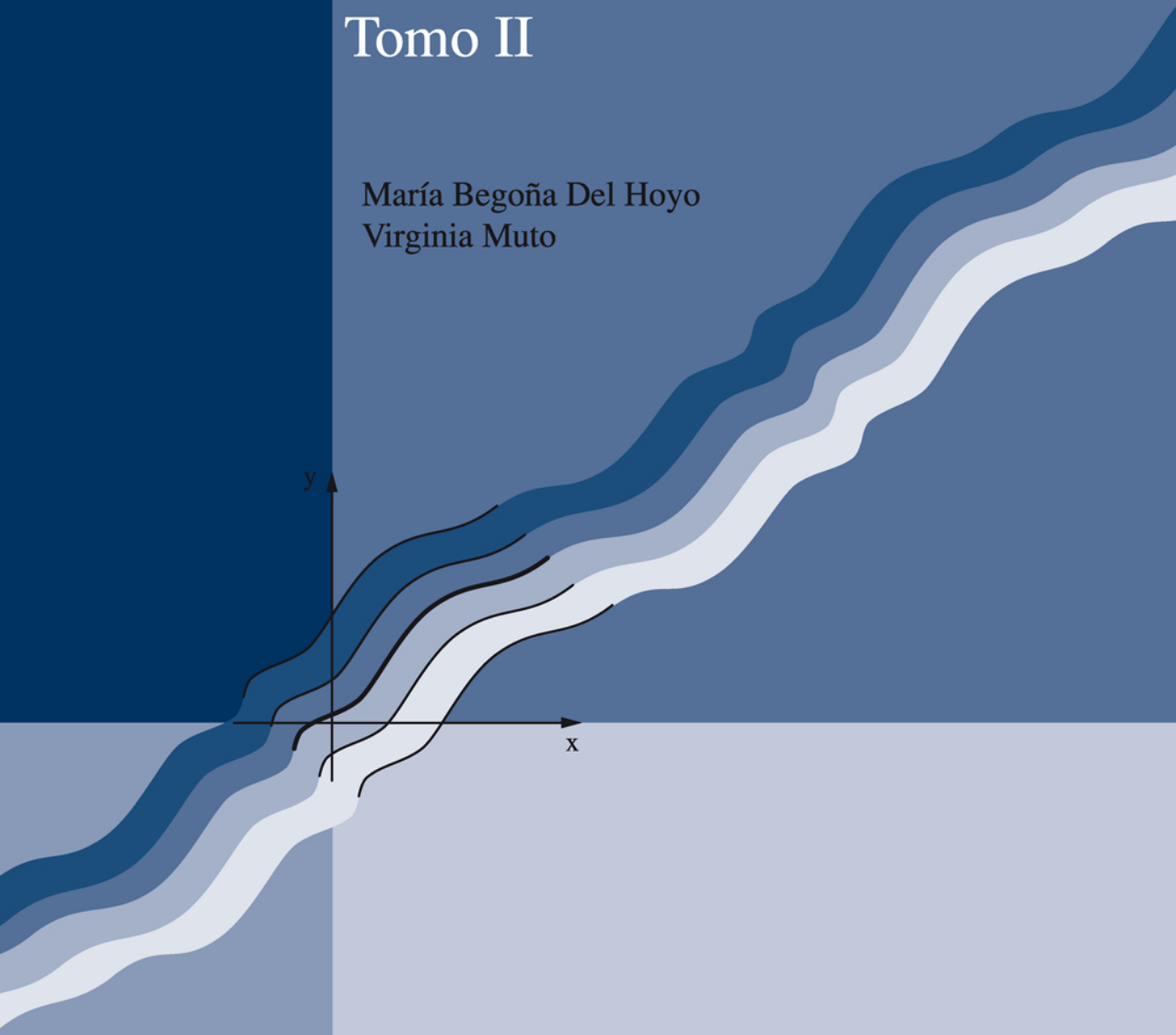


Fundamentos Matemáticos de la Ingeniería

Tomo II

María Begoña Del Hoyo
Virginia Muto



FUNDAMENTOS MATEMÁTICOS
DE LA INGENIERÍA

TOMO II

FUNDAMENTOS MATEMÁTICOS DE LA INGENIERÍA

TOMO II

María Begoña Del Hoyo y Virginia Muto

Departamento de Matemática Aplicada,
Estadística e Investigación Operativa /
Matematika Aplikatua eta Estatistika
eta Ikerkuntza Operatiboa Saila

Facultad de Ciencias / Zientzi Fakultatea
Universidad del País Vasco / Euskal Herriko Unibertsitatea

ormari ta zabal zazu



Universidad Euskal Herriko
del País Vasco Unibertsitatea
servicio editorial argitalpen zerbitzua

© Euskal Herriko Unibertsitateko Argitalpen Zerbitzua
Servicio Editorial de la Universidad del País Vasco

I.S.B.N.: 84-8373-370-6

Depósito Legal/Lege Gordailua: BI-2159-01

Impresión/Inprimatzea:

Servicio Editorial/Argitalpen Zerbitzua UPV/EHU

Prólogo

En el curso 1999-2000 comenzó a impartirse en la Facultad de Ciencias de la Universidad del País Vasco la titulación de Ingeniería Química.

Dentro de dicha titulación se nos asignó a las profesoras Virginia Muto y María Begoña del Hoyo, la docencia de la asignatura de "*Fundamentos Matemáticos para la Ingeniería*". Como el temario era amplio y el tiempo que teníamos para dedicar a nuestros alumnos era corto, decidimos preparar unos apuntes que les sirvieran de ayuda para "no perderse" por la asignatura y adentrarse así en el mundo de la Matemática que consideramos básica para poder desarrollar con éxito su formación en la Facultad.

Desafortunadamente nuestra experiencia nos dice que los alumnos que llegan a primero de Ingeniería Química carecen del hábito e incluso en muchos casos de la capacidad para el estudio y correcta interpretación de los textos usualmente disponibles en el mercado. La motivación principal que nos ha inducido a escribir este libro y probablemente nuestra aportación original esencial al mismo, ha sido la de producir un texto que pudiera jugar un papel intermedio entre el profesor-tutor y los textos ordinarios, por decirlo en otras palabras que proveyera al alumno con una especie de servicio de tutorización impresa, haciéndoles fácilmente digerible ideas, conceptos y métodos expresados de forma menos elaborada y compacta en otras obras.

Después de dos años de andar con hojas "para adelante y para atrás" hemos creído conveniente publicar estos apuntes en formato libro con el único objetivo de que para nuestros alumnos sea más fácil disponer de los apuntes completos de la asignatura desde el primer día de clase.

Para confeccionar dichos apuntes nos basamos en los siguientes libros:

- N. Piskunov: Cálculo Diferencial e Integral. Montaner y Simón, Barcelona, 1978.
- E. Marsden & A. J. Tromba: Cálculo Vectorial. Addison-Wesley Iberoamericana, 1987.
- L. Salas & E. Hille: Calculus - Cálculo de una y varias variables con geometría analítica. Reverté, 1995.
- B. P. Demidovich: Problemas y ejercicios de Análisis Matemático. Paraninfo, 1969.
- B. P. Demidovich: 5000 Problemas de Análisis Matemático. Paraninfo, Madrid, 1990.
- M. Spivak: Cálculos. Cálculo Infinitesimal. Reverté, 1970.
- J. Martínez Salas: Elementos de Matemáticas. Gráficas Andrés Martín, Valladolid, 1977.
- S. Lipschutz: Algebra Lineal. McGraw-Hill, 1992.
- M. Castellet & I. Llerena: Algebra Lineal y Geometría. Reverté, Barcelona, 1992.
- E. Tebar Flores: Problemas de Cálculo Infinitesimal. Tebar Flores, Albacete.
- A. Luzarraga: Problemas resueltos de Algebra Lineal. Romargraf, 1966.

y en otros apuntes que en nuestra experiencia docente de largos años habíamos ido acumulando.

Agradecemos cualquier sugerencia que los lectores nos hagan, para poder subsanar en el futuro defectos y errores en el trabajo que deben ser imputados a nuestra única responsabilidad. Siempre con el objetivo de conseguir un texto útil para nuestros alumnos de Ingeniería Química y si es posible, también, para alumnos de otras secciones.

Agradecemos la ayuda inestimable de César, sin la cual no habiésemos podido "arrancar" el 1 de Octubre de 1999, a nuestros maridos Jon y Alberto por sus primeras lecturas y críticas y a nuestros hijos Ane Miren, Mikel, Patxi y Ana que sobre todo el primer curso tuvieron que aguantar a unas madres que eran un manojo de nervios.

Leioa, Junio 2001
Virginia y Begoña

Índice General

I Tomo 1	13
-1 Introducción histórica	15
0 Conceptos Básicos.	17
0.1 Los números reales.	17
0.2 Sucesiones de números reales.	32
0.3 Funciones reales de variable real.	54
0.4 Límite de una función en un punto.	68
0.5 Infinitos e infinitésimos.	80
0.6 Continuidad de una función.	87
1 Derivación.	101
1.1 Derivada de una función.	101
1.2 Continuidad de las funciones de una variable.	105
1.3 Diferencial de una función.	117
1.4 Derivadas de orden superior.	120
1.5 Diferenciales de diversos ordenes.	124
2 Teoremas del valor medio. Aplicaciones.	129
2.1 Teoremas clásicos sobre derivación.	129
2.2 Regla de L'Hôpital.	134
2.3 Fórmula de Taylor. Resto de Lagrange.	137
2.4 Máximos y mínimos relativos y absolutos.	140
2.5 Trazado de curvas planas definidas de forma explícita.	153
3 Integral de Riemann.	169
3.1 Introducción a la integral de Riemann.	169
3.2 Propiedades de la integral definida.	171
3.3 Teorema del valor medio integral.	172
3.4 Teorema fundamental del cálculo integral.	173
3.5 Cambio de variable en la integral definida.	175
3.6 Aplicaciones de la integral definida.	176
3.6.1 Cálculo de áreas.	176
3.6.2 Cálculo de longitudes de arco de una curva.	178

3.6.3	Cálculo de volúmenes.	180
3.6.4	Cálculo del área de una superficie de revolución.	181
4	Técnicas de integración.	183
4.1	Función primitiva e integral indefinida.	183
4.2	Integrales inmediatas.	188
4.3	Integración por cambio de variable o sustitución.	189
4.4	Integración por partes.	192
4.5	Integración de fracciones racionales.	195
4.6	Integración de fracciones irracionales.	202
4.6.1	Diferenciales binomias.	203
4.6.2	Sustituciones de Euler.	206
4.7	Integración de funciones trigonométricas.	208
4.8	Integrales racionales de funciones trigonométricas.	213
4.9	Funciones hiperbólicas.	215
4.10	Método de Hermite.	216
5	Integrales impropias.	219
5.1	Definición de integral impropia.	219
5.2	Integrales extendidas a un intervalo infinito. (Integrales de 1ª especie).	219
5.3	Integrales de funciones no acotadas. (Integrales de 2ª especie).	221
5.4	Criterios de convergencia.	226
5.5	Criterio de convergencia de Cauchy.	231
6	Métodos Numéricos del Cálculo Integral.	235
6.1	Introducción.	235
6.2	Fórmula de los rectángulos.	237
6.3	Fórmula de los trapecios.	238
6.4	Fórmula de Simpson.	238
7	Series.	241
7.1	Conceptos previos.	241
7.2	Convergencia de una serie.	244
7.3	Criterios de convergencia.	248
7.3.1	Condición necesaria de convergencia de una serie.	248
7.3.2	Teoremas de comparación.	251
7.4	Series Alternadas.	265
7.5	Series de términos positivos y negativos.	267
7.5.1	Convergencia absoluta y condicional.	268
7.6	Sumación de Series.	273
7.6.1	Serie hipergeométrica.	273
7.6.2	Serie aritmético-geométrica.	273
7.6.3	Serie telescópica.	274
7.6.4	Series descomponibles en factores simples.	275

ÍNDICE GENERAL

7.6.5	Series descomponibles mediante el número e .	276
7.7	Series de Funciones.	276
7.8	Series de potencias. Intervalo de convergencia.	280
7.8.1	Series de Taylor y de MacLaurin.	284
7.8.2	Serie binómica.	291
II Tomo 2		309
8	Algebra Matricial.	311
8.1	Matrices. Operaciones con matrices.	311
8.2	Aplicaciones lineales.	317
8.3	Matriz asociada a una aplicación lineal.	323
8.4	Suma, producto y composición de aplicaciones lineales.	328
8.5	Determinante de una matriz $n \times n$. Propiedades.	329
8.6	Cálculo de determinantes.	335
8.7	Tipos de matrices.	337
8.8	Normas de vectores y de matrices.	345
8.9	Autovalores y autovectores.	349
8.10	Diagonalización de matrices.	357
9	La geometría del espacio euclídeo.	363
9.1	Vectores en el espacio tridimensional.	363
9.2	Producto escalar.	370
9.3	Producto vectorial.	374
9.4	Coordenadas esféricas y cilíndricas.	381
9.5	Espacio euclídeo de dimensión n .	385
10	Diferenciación.	389
10.1	La geometría de las funciones con valores reales.	389
10.2	Límites y continuidad.	398
10.3	Diferenciación.	409
10.4	Propiedades de la derivada.	418
10.5	Gradientes y derivadas direccionales.	424
10.6	Derivadas parciales iteradas.	433
10.7	Trayectorias y velocidad. Longitud de arco.	438
11	Derivadas de orden superior: máximos y mínimos.	445
11.1	Teorema de Taylor.	445
11.2	Extremos de funciones con valores reales.	448
11.3	Extremos con restricciones y multiplicadores de Lagrange.	460
11.4	Métodos numéricos de optimización con y sin restricciones.	466

12 Integración.	471
12.1 Introducción.	471
12.2 La integral doble sobre un rectángulo.	475
12.3 La integral doble sobre regiones más generales.	485
12.4 Cambio del orden de integración.	490
12.5 La integral triple.	497
12.6 Cambio de variables en la integral múltiple.	507
12.6.1 La geometría de las funciones de \mathbb{R}^2 a \mathbb{R}^2 .	507
12.6.2 Cambio de variables en la integral doble.	511
12.6.3 Cambio de variables en la integral triple.	522
12.7 Algunas aplicaciones de las integrales múltiples.	527
13 Integrales sobre trayectorias y superficies.	535
13.1 Campos vectoriales.	535
13.1.1 Divergencia y rotacional de un campo vectorial.	540
13.1.2 Cálculo diferencial vectorial.	545
13.2 La integral de trayectoria.	548
13.3 La integral de línea.	551
13.4 Superficies parametrizadas.	568
13.5 Integrales de funciones escalares sobre superficies.	579
13.6 Integrales de superficie de funciones vectoriales.	585
13.7 Teorema de Green.	594
13.8 Teorema de Stokes.	601
13.9 Teorema de Gauss.	609

Tomo 2

Capítulo 8

Algebra Matricial.

8.1 Matrices. Operaciones con matrices.

Definición. Una **matriz** rectangular de orden $m \times n$ sobre un cuerpo K es una disposición rectangular de $m \cdot n$ números encerrada entre paréntesis, distribuidos en m filas y n columnas. A los números que aparecen en una matriz se les llama **coeficientes** de la matriz. A cada uno de estos elementos se le atribuyen dos subíndices que indican el primero la fila y el segundo la columna de la matriz, esto es, la posición en la que está dicho elemento.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij}), \quad a_{ij} \in K, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n.$$

Ejemplos concretos de matrices 2×2 , 2×4 y 3×3 son los siguientes:

$$\begin{pmatrix} 2 & 4 \\ 3 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 2 & 3 \\ 6 & 5 & 4 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 5 & 2 \\ 3 & 7 & 2 \\ 6 & 5 & 4 \end{pmatrix}.$$

A la primera y a la tercera se las llama **matrices cuadradas**: tienen el mismo número de filas y columnas.

Definición. Se llama **diagonal principal** de una matriz cuadrada $n \times n$ a todos los elementos que se encuentran en la línea que va desde el vértice superior de la izquierda (a_{11}) al inferior de la derecha (a_{nn}). A la otra diagonal se le llama **secundaria** (desde a_{1n} a a_{n1}).

Definición. Una **matriz fila** es una matriz de orden $1 \times n$

$$A = (a_{11} \ a_{12} \ \dots \ a_{1n}).$$

Definición. Una **matriz columna** es una matriz de orden $m \times 1$

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}.$$

Definición. Dada una matriz A , se llama **matriz opuesta**, y se representa por $-A$, a aquella matriz tal que sus elementos son los simétricos (opuestos en el caso $K = \mathbb{R}$) de los elementos de la matriz A .

Por ejemplo

$$A = \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix}, \quad -A = \begin{pmatrix} -2 & -4 \\ -3 & -6 \end{pmatrix}.$$

Definición. Una matriz A de orden $m \times n$ es **nula** si todos los elementos a_{ij} son nulos. Se suele denotar con el símbolo 0 (cero).

Ejemplo. Las siguientes matrices nulas son de orden 3×2 y 3×3 , respectivamente:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definición. Una matriz A cuadrada $n \times n$ es **diagonal** si todos los elementos que no pertenezcan a la diagonal principal son nulos.

Por ejemplo, las siguientes matrices 3×3 son diagonales:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Definición. Se dice que dos matrices A y B son **iguales** si son del mismo tamaño $m \times n$ y si $a_{ij} = b_{ij}$ para cada $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, n$.

Pasamos ahora a estudiar las operaciones con matrices. Sea $M_{m \times n}$ el conjunto de matrices de orden $m \times n$ sobre el cuerpo K .

Definición. Si A y B son matrices ambas $m \times n$, $A, B \in M_{m \times n}$, entonces la **suma** de A y B , denotada por $A + B$, es la matriz $m \times n$ cuyos elementos son $a_{ij} + b_{ij}$, para cada $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, n$. Para sumar dos matrices, las dos tienen que ser del mismo orden.

De esta manera, se ha definido una ley de composición interna:

$$\begin{matrix} M_{m \times n} \times M_{m \times n} \\ (A, B) \end{matrix} \Rightarrow \begin{matrix} M_{m \times n} \\ A + B \end{matrix}.$$

8.1. MATRICES. OPERACIONES CON MATRICES.

Ejemplo. Suma de dos matrices 2×3 :

$$\begin{pmatrix} 2 & 4 & 0 \\ 0 & 3 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 6 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 3 \\ 6 & 7 & 4 \end{pmatrix}.$$

Definición. Si A es una matriz $m \times n$ y λ es un escalar, $\lambda \in K$, entonces el **producto escalar** de λ por A , denotado como λA , es la matriz $m \times n$ cuyos elementos son λa_{ij} , para cada $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, n$; es decir, la matriz que resulta al multiplicar cada elemento de la matriz A por λ .

Ejemplo de producto escalar de $\lambda = 3$ y A matriz 3×3 es el siguiente:

$$\lambda A = 3 \begin{pmatrix} 2 & -2 & 0 \\ 0 & 3 & 3 \\ 2 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 6 & -6 & 0 \\ 0 & 9 & 9 \\ 6 & -3 & 9 \end{pmatrix}.$$

Si $A \in M_{m \times n}$, la matriz $\lambda A \in M_{m \times n}$, luego acabamos de definir una ley de composición externa:

$$K \times M_{m \times n} \quad \rightarrow \quad M_{m \times n} \\ (\lambda, A) \quad \mapsto \quad \lambda A.$$

Podemos ahora enumerar las siguientes **propiedades generales** de la adición y de la multiplicación escalar matricial. Estas propiedades son suficientes para clasificar el conjunto $M_{m \times n}$ de todas las matrices $m \times n$ con elementos en un cuerpo K (en nuestro caso $K = \mathbb{R}$ ó $K = \mathbb{C}$), como un espacio vectorial sobre el cuerpo K .

1. Propiedad asociativa de la suma: $\forall A, B, C \in M_{m \times n}$

$$(A + B) + C = A + (B + C);$$

2. Elemento neutro de la suma: $\forall A \in M_{m \times n}$, existe la matriz nula $0 \in M_{m \times n}$ tal que

$$A + 0 = 0 + A = A;$$

3. Elemento simétrico: $\forall A \in M_{m \times n}$, existe la matriz opuesta $-A \in M_{m \times n}$ tal que

$$A + (-A) = (-A) + A = 0;$$

4. Propiedad conmutativa de la suma: $\forall A, B \in M_{m \times n}$

$$A + B = B + A.$$

Con estas primeras cuatro propiedades tenemos que $(M_{m \times n}, +)$ es un grupo conmutativo.

5. Propiedad distributiva del producto escalar con respecto a la suma de matrices:
 $\forall \lambda \in K, \forall A, B \in M_{m \times n}$

$$\lambda(A + B) = \lambda A + \lambda B;$$

6. Propiedad distributiva del producto escalar con respecto a la suma escalar:
 $\forall \lambda, \mu \in K, \forall A \in M_{m \times n}$

$$(\lambda + \mu) A = \lambda A + \mu A;$$

7. Propiedad asociativa del producto escalar: $\forall \lambda, \mu \in K, \forall A \in M_{m \times n}$

$$\lambda(\mu A) = (\lambda \mu) A;$$

8. Elemento neutro del producto escalar: $\forall A \in M_{m \times n}$, existe e_1 elemento neutro para el producto en el cuerpo K , tal que

$$e_1 \cdot A = A;$$

Todas estas propiedades se cumplen por ser K un cuerpo.

Con estas últimas cuatro propiedades tenemos que $(M_{m \times n}, +, \cdot)$ es un espacio vectorial sobre el cuerpo K .

Consideremos ahora las $m \cdot n$ matrices M_{ij} , de orden $m \times n$ con todos los elementos nulos menos el correspondiente a los subíndices de la matriz, que vale uno. Por ejemplo

$$M_{11} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad M_{2n} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad M_{mn} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Es fácil verificar que dada una matriz cualquiera A de orden $m \times n$ es posible escribir la matriz A como suma de las $m \cdot n$ matrices M_{ij}

$$A = a_{11} M_{11} + a_{12} M_{12} + a_{13} M_{13} + \dots + a_{mn} M_{mn}.$$

Esto quiere decir que $\{M_{11}, M_{12}, \dots, M_{1n}, \dots, M_{mn}\}$ es un sistema de generadores del espacio vectorial $(M_{m \times n}, +, \cdot)$. Además, son linealmente independientes por ser

$$\alpha_{11} M_{11} + \alpha_{12} M_{12} + \alpha_{13} M_{13} + \dots + \alpha_{mn} M_{mn} = 0$$

si y sólo si todos los escalares α_{ij} son nulos. Luego las matrices $\{M_{11}, M_{12}, \dots, M_{mn}\}$ forman una base del espacio vectorial de las matrices $M_{m \times n}$ definidas sobre el cuerpo K . Entonces, la dimensión de este espacio será $m \cdot n$.

8.1. MATRICES. OPERACIONES CON MATRICES.

Definición. Sean A una matriz $m \times n$ y B una matriz $n \times p$, definidas sobre un cuerpo K . El **producto matricial** de A y B , denotado por $C = AB$, es una matriz $m \times p$, cuyos elementos c_{ij} están dados por

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}$$

para cada $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, p$ (normalmente conocido como producto filas por columnas). Es importante notar que para poder multiplicar dos matrices A y B el número de columnas de la matriz A ha de coincidir con el número de filas de la matriz B .

Ejemplo. Halla el producto matricial de las dos matrices A , 2×3 , y B , 3×4 .

$$AB = \begin{pmatrix} 2 & 4 & 0 \\ 0 & 3 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 2 & -1 & 1 \\ 1 & 2 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 12 & 2 & 4 \\ 3 & 12 & -6 & -3 \end{pmatrix}.$$

Definición. Una matriz diagonal $n \times n$ en K es **matriz unidad** o **identidad** de orden n , si $AI = IA = A$. En el caso $K = \mathbb{R}$, todos los elementos de la diagonal principal son unos, y se representa por I_n o simplemente I .

Ejemplo. La matriz identidad de orden 3 es

$$I = I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matriz identidad de orden n conmuta con una matriz A de orden n , es decir, el orden de la multiplicación no importa. Por otra parte, la propiedad conmutativa, $AB = BA$, no es generalmente cierta para la multiplicación matricial. Por ejemplo, si consideremos las matrices A y B del ejemplo anterior, mientras es posible calcular AB , no está definido el producto BA . Esto es, para que sea posible hallar los dos productos AB y BA , antes de todo las matrices A y B han de ser matrices cuadradas. Si consideremos las siguientes dos matrices 2×2 , tendremos

$$AB = \begin{pmatrix} 2 & 4 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} = \begin{pmatrix} -2 & -4 \\ -3 & -6 \end{pmatrix},$$

$$BA = \begin{pmatrix} 1 & 2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 10 \\ -2 & -10 \end{pmatrix}.$$

Algunas de las **propiedades** relacionadas con la multiplicación de matrices, que sí se satisfacen, se presentan a continuación:

1. Propiedad asociativa del producto: $\forall A \in M_{m \times n}, B \in M_{n \times p}, C \in M_{p \times q}$

$$A(BC) = (AB)C;$$

2. Propiedad distributiva del producto matricial con respecto a la suma de matrices:
 $\forall A \in M_{m \times n}, B, C \in M_{n \times p}$

$$A(B + C) = AB + AC;$$

$$\forall A, B \in M_{m \times n}, C \in M_{n \times p}$$

$$(A + B)C = AC + BC;$$

3. Propiedad distributiva del producto escalar con respecto al producto matricial:
 $\forall \lambda \in K, \forall A, B \in M_{m \times n}$

$$\lambda(AB) = (\lambda A)B = A(\lambda B);$$

4. Elemento neutro del producto: $\forall A \in M_{m \times n}$, existen $I_m \in M_{m \times m}$ y $I_n \in M_{n \times n}$ elementos neutros para el producto. tal que

$$I_m A = A, \quad A I_n = A.$$

Ejercicios.

- 1) Determina dos matrices X e Y que verifiquen el sistema

$$\begin{cases} 3X + 4Y = \begin{pmatrix} 5 & 2 \\ -12 & 9 \end{pmatrix} \\ -2X + 3Y = \begin{pmatrix} 8 & -7 \\ -9 & -6 \end{pmatrix}. \end{cases}$$

- 2) Calcula

a) $\begin{pmatrix} 1 & -1 & 6 & -1 \\ 2 & 0 & 2 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 5 & 1 & 3 \\ 3 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix};$

b) $\begin{pmatrix} 1 & 0 & 2 \\ 5 & 1 & 3 \\ 3 & 0 & 1 \\ 2 & 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 6 & -1 \\ 2 & 0 & 2 & 0 \\ 1 & 2 & 1 & 2 \end{pmatrix}.$

- 3) Dadas las matrices $A = \begin{pmatrix} 4 & 3 & 2 \\ 7 & 6 & 3 \\ 2 & 9 & 8 \end{pmatrix}$ y $B = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 3 & 2 \\ 6 & 7 & 4 \end{pmatrix}$, calcula

a) $A + B,$

8.2. APLICACIONES LINEALES.

- b) $A \cdot B$,
c) $A^2 \cdot B$,
d) $B \cdot A$,
e) $2A + 3B$.
- 4) Dada la matriz $A = \begin{pmatrix} 1 & -2 \\ -3 & 4 \end{pmatrix}$, halla todas las matrices que conmuten con ella, es decir tales que $A \cdot B = B \cdot A$.
- 5) Dada la matriz $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, halla $B = A + A^2 + \dots + A^n$.
- 6) Dada la matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & m \end{pmatrix}$, halla "m" para que existan matrices no nulas B tales que $A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.
- 7) Calcula la potencia n -ésima de las matrices

a) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, b) $B = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

- 8) Dada la matriz $A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$, calcula A^2, A^3, \dots, A^n .

8.2 Aplicaciones lineales.

Definición. Sean $(E, +, \cdot)$ y $(F, +, \cdot)$ dos espacios vectoriales sobre el mismo cuerpo K . Una aplicación $f: E \rightarrow F$ se dice que es una **aplicación lineal** (u homomorfismo) entre los espacios vectoriales E y F si se cumple:

1. $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$, $\forall \vec{x}, \vec{y} \in E$
2. $f(\alpha \vec{x}) = \alpha f(\vec{x})$, $\forall \alpha \in K, \forall \vec{x} \in E$.

Teorema 8.1 La condición necesaria y suficiente para que una aplicación entre dos espacios vectoriales E y F definidos sobre el mismo cuerpo K sea lineal es que

$$\forall \alpha, \beta \in K, \quad \forall \vec{x}, \vec{y} \in E \quad f(\alpha \vec{x} + \beta \vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}).$$

Demostración: Para simplificar sea $K = \mathbb{R}$.

$$\xrightarrow{N} \quad \forall \alpha, \beta \in K, \quad \forall \vec{x}, \vec{y} \in E$$

$$f(\alpha \vec{x} + \beta \vec{y}) = f(\alpha \vec{x}) + f(\beta \vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y})$$

habiendo usado las dos reglas que definen una aplicación lineal.

$$\xleftarrow{S} \quad \text{Ahora partimos de que se cumple}$$

$$\forall \alpha, \beta \in K, \quad \forall \vec{x}, \vec{y} \in E \quad f(\alpha \vec{x} + \beta \vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}),$$

si tomamos $\alpha = \beta = 1$ obtenemos directamente la primera regla. Escogiendo $\beta = 0$ obtenemos la segunda. c.q.d.

Sabemos que todo cuerpo K puede ser considerado como un espacio vectorial sobre sí mismo. Se puede entonces dar la siguiente:

Definición. Si la aplicación $f : E \rightarrow K$ es lineal, siendo E un espacio vectorial sobre K , se dice que f es una **forma lineal** en (o sobre) E .

Como caso particular tenemos que la forma lineal en K es una aplicación lineal $f : K \rightarrow K$.

Pasamos ahora a dar una clasificación de las aplicaciones lineales:

Sea $f : E \rightarrow F$ una aplicación lineal, con E y F espacios vectoriales sobre K .

- Si f es inyectiva: **monomorfismo**;
- si f es suprayectiva: **epimorfismo**;
- si f es biyectiva: **isomorfismo**;
- si $E = F$: **endomorfismo en E** ;
- si $E = F$ y f es biyectiva: **automorfismo**.

Definición. Sean E, F espacios vectoriales sobre K , y $f : E \rightarrow F$ una aplicación lineal: el **núcleo** de la aplicación f es el conjunto de vectores de E cuya imagen es el vector cero de F :

$$\text{Ker}(f) = \{ \vec{x} \in E / f(\vec{x}) = \vec{0}_F \}.$$

Claramente el núcleo de una aplicación f es un subespacio vectorial de E .

- $\text{Ker}(f) \subset E$;
- dado que $f(\vec{0}_E) = \vec{0}_F$, luego $\vec{0}_E \in \text{Ker}(f)$, es decir, $\text{Ker}(f) \neq \emptyset$.
- $\forall \alpha, \beta \in K, \quad \forall \vec{x}, \vec{y} \in \text{Ker}(f)$ resulta que

$$f(\alpha \vec{x} + \beta \vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}) = \alpha \vec{0}_F + \beta \vec{0}_F = \vec{0}_F$$

es decir, también $\alpha \vec{x} + \beta \vec{y} \in \text{Ker}(f)$.

8.2. APLICACIONES LINEALES.

Definición. Sean E, F espacios vectoriales sobre K , y $f : E \rightarrow F$ una aplicación lineal; la **imagen** de la aplicación f es el conjunto de vectores de F que son imagen de un vector \vec{x} de E :

$$\text{Img}(f) = f(E) = \{\vec{y} \in F / \exists \vec{x} \in E \text{ tal que } f(\vec{x}) = \vec{y}\}.$$

Claramente la imagen de una aplicación f es un subespacio vectorial de F .

- $\text{Img}(f) \subset F$: por definición;
- dado que $f(\vec{0}_E) = \vec{0}_F$, luego $\vec{0}_F \in \text{Img}(f)$, es decir, $\text{Img}(f) \neq \emptyset$.
- $\forall \alpha, \beta \in K, \forall \vec{y}_1, \vec{y}_2 \in \text{Img}(f)$ existen $\vec{x}_1, \vec{x}_2 \in E$ para los cuales resulta que

$$\alpha \vec{y}_1 + \beta \vec{y}_2 = \alpha f(\vec{x}_1) + \beta f(\vec{x}_2) = f(\alpha \vec{x}_1 + \beta \vec{x}_2)$$

es decir, también $\alpha \vec{y}_1 + \beta \vec{y}_2 \in \text{Img}(f)$.

Pasamos ahora a estudiar algunas **propiedades** de las aplicaciones lineales:

- (1) Toda aplicación lineal transforma el vector nulo del espacio vectorial E , $\vec{0}_E$, en el vector nulo del espacio vectorial F , $\vec{0}_F$.

Demostración:

$$\forall \vec{x} \in E, \quad \vec{x} = \vec{x} + \vec{0}_E \quad \iff \quad f(\vec{x}) = f(\vec{x}) + f(\vec{0}_E)$$

$$\forall \vec{x} \in E, \quad f(\vec{x}) \in F \quad \iff \quad f(\vec{x}) = f(\vec{x}) + \vec{0}_F$$

entonces $f(\vec{0}_E) = \vec{0}_F$. c.q.d.

- (2) Si f es una aplicación lineal, la imagen del simétrico (u opuesto) de un vector coincide con el simétrico de su imagen.

Demostración:

$$\forall \vec{x} \in E, \quad \vec{x} + (-\vec{x}) = \vec{0}_E \quad \iff \quad f(\vec{x} + (-\vec{x})) = f(\vec{x}) + f(-\vec{x}) = f(\vec{0}_E) = \vec{0}_F$$

$$\forall \vec{x} \in E, \quad f(\vec{x}) \in F \quad \iff \quad f(\vec{x}) + (-f(\vec{x})) = \vec{0}_F$$

entonces $f(-\vec{x}) = -f(\vec{x})$. c.q.d.

- (3) Toda aplicación lineal transforma un sistema de vectores linealmente dependientes del espacio E en un sistema de vectores linealmente dependientes de F .

Demostración:

Sean $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ n vectores linealmente dependientes, pertenecientes a E . Entonces, existe una combinación lineal con no todos los α_i nulos tal que

$$\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n = \vec{0}_E$$

Como f es una aplicación lineal tendremos

$$f(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n) = \alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2) + \dots + \alpha_n f(\vec{u}_n) = f(\vec{0}_E) = \vec{0}_F$$

y dado que no todos los α_i son nulos significa que $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$ son vectores de F linealmente dependientes.

c.q.d.

- (4) Toda aplicación lineal transforma un sistema de generadores del espacio E en un sistema de generadores del subespacio imagen $\text{Img}(f) = f(E) \subset F$.

Demostración:

Sea $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ un sistema de generadores de E . Sea $\vec{y} \in f(E)$, entonces existe $\vec{x} \in E$ tal que $f(\vec{x}) = \vec{y}$; como $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ es un sistema de generadores de E y $\vec{x} \in E$ existirán $\alpha_1, \alpha_2, \dots, \alpha_n \in K$ de manera que

$$\vec{x} = \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n$$

y como f es una aplicación lineal tendremos

$$\vec{y} = f(\vec{x}) = f(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n) = \alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2) + \dots + \alpha_n f(\vec{u}_n).$$

es decir, \vec{y} es combinación lineal de los elementos $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$, los cuales constituyen un sistema de generadores de $\text{Img}(f) = f(E) \subset F$.

c.q.d.

Claramente si la aplicación es suprayectiva entonces $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$ es un sistema de generadores de F .

- (5) Toda aplicación lineal inyectiva transforma un sistema de vectores linealmente independientes del espacio E en un sistema de vectores linealmente independientes del espacio F .

Demostración:

Sea $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ un sistema de vectores linealmente independientes del espacio E . Sea

$$\alpha_1 f(\vec{u}_1) + \alpha_2 f(\vec{u}_2) + \dots + \alpha_n f(\vec{u}_n) = \vec{0}_F = f(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n)$$

dado que f es lineal. Esto quiere decir que $\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n \in \text{Ker}(f)$, y entonces

$$\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_n \vec{u}_n = \vec{0}_E$$

pero como los vectores u_i son linealmente independientes en E , será $\alpha_i = 0$ para todo $i = 1, 2, \dots, n$. Entonces, los elementos $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$, son vectores linealmente independientes de F .

c.q.d.

8.2. APLICACIONES LINEALES.

- (6) Si f es aplicación lineal biyectiva transforma una base del espacio E en una base del espacio F .

Demostración:

Sea $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ una base del espacio E .

Como f es inyectiva, $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$ son vectores linealmente independientes; como f es suprayectiva, $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$ es un sistema de generadores de F , en conclusión $\{f(\vec{u}_1), f(\vec{u}_2), \dots, f(\vec{u}_n)\}$ es una base de F .

c.q.d.

Teorema 8.2 Sean E, F espacios vectoriales sobre K , y $f : E \rightarrow F$ una aplicación lineal. La aplicación lineal f es un monomorfismo si y sólo si $\text{Ker}(f) = \{\vec{0}_F\}$.

Demostración:

$\xrightarrow{N.}$ si $\vec{x} \in \text{Ker}(f)$, entonces $f(\vec{x}) = \vec{0}_F = f(\vec{0}_E)$ y como f es inyectiva, se sigue que

$$\vec{x} = \vec{0}_E. \quad \text{Luego, } \text{Ker}(f) = \{\vec{0}_F\}.$$

$\xleftarrow{S.}$ si $f(\vec{x}) = f(\vec{y})$ entonces $f(\vec{x}) + f(-\vec{y}) = \vec{0}_F$, lo que quiere decir que

$$\vec{x} + (-\vec{y}) \in \text{Ker}(f), \quad \text{es decir } \vec{x} + (-\vec{y}) = \vec{0}_E,$$

o sea $\vec{x} = \vec{y}$ o lo que es lo mismo f es inyectiva.

c.q.d.

Teorema 8.3 Sean E y F espacios vectoriales sobre K y $f : E \rightarrow F$ una aplicación lineal. Se cumple

$$\dim E = \dim \text{Ker}(f) + \dim \text{Img}(f).$$

Demostración:

Supongamos que $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\}$ sea una base de $\text{Ker}(f)$, luego $\dim \text{Ker}(f) = p$. También supongamos que $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ sea una base de $\text{Img}(f)$, luego $\dim \text{Img}(f) = n$. Ahora resulta que:

para $\vec{v}_1 \in F$ existe $\vec{u}_{p+1} \in E$ tal que $f(\vec{u}_{p+1}) = \vec{v}_1$,

para $\vec{v}_2 \in F$ existe $\vec{u}_{p+2} \in E$ tal que $f(\vec{u}_{p+2}) = \vec{v}_2$,

es decir, para cada $\vec{v}_i \in F$ existe $\vec{u}_{p+i} \in E$ tal que $f(\vec{u}_{p+i}) = \vec{v}_i$, con $i = 1, 2, \dots, n$.

Veamos ahora que $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p, \vec{u}_{p+1}, \dots, \vec{u}_{p+n}\}$ es una base de E . Para todo elemento $\vec{x} \in E$ tendremos que $f(\vec{x}) \in \text{Img}(f)$, entonces por ser f aplicación lineal

$$f(\vec{x}) = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_n \vec{v}_n = \alpha_1 f(\vec{u}_{p+1}) + \alpha_2 f(\vec{u}_{p+2}) + \dots + \alpha_n f(\vec{u}_{p+n})$$

quiere decir que

$$f(\vec{x}) = f(\alpha_1 \vec{u}_{p+1} + \alpha_2 \vec{u}_{p+2} + \dots + \alpha_n \vec{u}_{p+n}) = f(\vec{x}_1).$$

Entonces,

$$f(\vec{x}) - f(\vec{x}_1) = \vec{0}_F = f(\vec{x}) + f(-\vec{x}_1) = f(\vec{x} + (-\vec{x}_1)) \implies \vec{x} + (-\vec{x}_1) \in \text{Ker}(f)$$

luego $\vec{x} + (-\vec{x}_1)$ se puede expresar en función de los vectores de B_1 :

$$\vec{x} + (-\vec{x}_1) = \beta_1 \vec{u}_1 + \beta_2 \vec{u}_2 + \cdots + \beta_p \vec{u}_p$$

que nos lleva a

$$\begin{aligned} \vec{x} &= \beta_1 \vec{u}_1 + \beta_2 \vec{u}_2 + \cdots + \beta_p \vec{u}_p + \vec{x}_1 = \\ &= \beta_1 \vec{u}_1 + \beta_2 \vec{u}_2 + \cdots + \beta_p \vec{u}_p + \alpha_1 \vec{u}_{p+1} + \alpha_2 \vec{u}_{p+2} + \cdots + \alpha_n \vec{u}_{p+n}, \end{aligned}$$

es decir, $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p, \vec{u}_{p+1}, \dots, \vec{u}_{p+n}\}$ es un sistema de generadores de E . Nos queda mirar que son linealmente independientes. Por eso, sea

$$\gamma_1 \vec{u}_1 + \gamma_2 \vec{u}_2 + \cdots + \gamma_p \vec{u}_p + \gamma_{p+1} \vec{u}_{p+1} + \gamma_{p+2} \vec{u}_{p+2} + \cdots + \gamma_{p+n} \vec{u}_{p+n} = \vec{0}_E \quad (8.1)$$

como f es lineal:

$$\gamma_1 f(\vec{u}_1) + \gamma_2 f(\vec{u}_2) + \cdots + \gamma_p f(\vec{u}_p) + \gamma_{p+1} f(\vec{u}_{p+1}) + \gamma_{p+2} f(\vec{u}_{p+2}) + \cdots + \gamma_{p+n} f(\vec{u}_{p+n}) = \vec{0}_F.$$

Por ser $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p \in \text{Ker}(f)$ los primeros p términos dan el $\vec{0}_F$ y nos queda

$$\gamma_{p+1} f(\vec{u}_{p+1}) + \gamma_{p+2} f(\vec{u}_{p+2}) + \cdots + \gamma_{p+n} f(\vec{u}_{p+n}) = \vec{0}_F = \gamma_{p+1} \vec{v}_1 + \gamma_{p+2} \vec{v}_2 + \cdots + \gamma_{p+n} \vec{v}_n.$$

Como B_2 es una base de $\text{Im}(f)$ sigue

$$\gamma_{p+1} = \gamma_{p+2} = \cdots = \gamma_{p+n} = 0$$

luego en (8.1) quedamos con

$$\gamma_1 \vec{u}_1 + \gamma_2 \vec{u}_2 + \cdots + \gamma_p \vec{u}_p = \vec{0}_E$$

y como B_1 es una base de $\text{Ker}(f)$ sigue

$$\gamma_1 = \gamma_2 = \cdots = \gamma_p = 0.$$

Es decir, partiendo de la ecuación (8.1), hemos llegado a que todos los escalares γ_i , con $i = 1, 2, \dots, p+n$, son nulos. Luego los vectores $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p, \vec{u}_{p+1}, \dots, \vec{u}_{p+n}\}$ además de ser un sistema de generadores de E , son linealmente independientes, es decir constituyen una base de E . Por lo tanto: $\dim E = p + n$.

c.q.d.

Como consecuencia de los teoremas 8.2 y 8.3 resulta que si tenemos un endomorfismo inyectivo $\dim \text{Ker}(f) = 0$, es decir $\dim E = \dim \text{Im}(f)$, o sea la aplicación es también suprayectiva. Siendo la aplicación lineal f biyectiva resulta ser un automorfismo.

8.3. MATRIZ ASOCIADA A UNA APLICACIÓN LINEAL.

Ejercicios.

- 1) Sean las aplicaciones $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f(x, y) = (y, x)$ y $g(x, y) = (3x, 3x + y)$. Demuestra que son lineales. Calcula los núcleos $\text{Ker}(f)$ y clasifica las aplicaciones.
- 2) Demuestra que la aplicación $f : (\mathbb{R}^2, +, \cdot) \rightarrow (\mathbb{R}^3, +, \cdot)$, $f(x, y) = (x + y, x, x - y)$, es una aplicación lineal inyectiva no suprayectiva.
- 3) ¿Las aplicaciones $f, g : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^4$ y $g(x) = 3x$, son formas lineales?
- 4) Demuestra que las aplicaciones siguientes son lineales. Calcula $\text{Ker}(f)$, $\dim \text{Ker}(f)$, $\dim \text{Im}(f)$ y clasificalas:
 - a) $f : (\mathbb{R}^2, +, \cdot) \rightarrow (\mathbb{R}^2, +, \cdot) / f(x, y) = (2x - y, x + y)$,
 - b) $f : (\mathbb{R}^3, +, \cdot) \rightarrow (\mathbb{R}^3, +, \cdot) / f(x, y, z) = (y - z, 4z - x, y - z)$,
 - c) $f : (\mathbb{R}^3, +, \cdot) \rightarrow (\mathbb{R}^2, +, \cdot) / f(x, y, z) = (x + y, y - z)$,
 - d) $f : (\mathbb{R}^3, +, \cdot) \rightarrow (\mathbb{R}^3, +, \cdot) / f(x, y, z) = (0, x + y, -z)$,
 - e) $f : (\mathbb{R}^2, +, \cdot) \rightarrow (\mathbb{R}^4, +, \cdot) / f(x, y) = (x, y, 0, 0)$.

8.3 Matriz asociada a una aplicación lineal.

Consideremos dos espacios vectoriales $(E, +, \cdot)$ y $(F, +, \cdot)$ definidos sobre un cuerpo K . Sean $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ una base de E y $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ una base de F . Sea $f : E \rightarrow F$ una aplicación lineal. Luego podremos escribir

$$\begin{aligned} f(\vec{u}_1) &= a_{11} \vec{v}_1 + a_{21} \vec{v}_2 + \dots + a_{m1} \vec{v}_m \\ f(\vec{u}_2) &= a_{12} \vec{v}_1 + a_{22} \vec{v}_2 + \dots + a_{m2} \vec{v}_m \\ &\vdots \\ f(\vec{u}_n) &= a_{1n} \vec{v}_1 + a_{2n} \vec{v}_2 + \dots + a_{mn} \vec{v}_m \end{aligned}$$

Definición. Se llama **matriz asociada** a la aplicación lineal f en las bases B_1 y B_2 a la matriz

$$M(f)_{B_1, B_2} = A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \quad (8.2)$$

Nótese que las columnas de la matriz asociada a la aplicación f están formadas por las componentes de las imágenes de la base B_1 en función de la base B_2 .

Ejemplo. Dada la aplicación lineal $f : (\mathbb{R}^3, +, \cdot) \rightarrow (\mathbb{R}^3, +, \cdot)$ tal que

$$f : (x_1, x_2, x_3) \rightsquigarrow (x_1 + x_2, x_2, 0)$$

la matriz asociada a esta aplicación lineal cuando en el espacio inicial y en el final se consideran las bases canónicas se obtiene de la siguiente manera:

$$\begin{aligned} f(\vec{e}_1) &= f(1, 0, 0) = (1, 0, 0) = 1 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \\ f(\vec{e}_2) &= f(0, 1, 0) = (1, 1, 0) = 1 \cdot \vec{e}_1 + 1 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \\ f(\vec{e}_3) &= f(0, 0, 1) = (0, 0, 0) = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \end{aligned}$$

es decir,

$$M(f)_{C,C} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Si se consideran en el espacio inicial la base $B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, donde $\vec{u}_1 = (2, 0, 0)$, $\vec{u}_2 = (-1, 3, 0)$ y $\vec{u}_3 = (0, 0, 5)$, y en el espacio final la base canónica se obtiene la siguiente matriz asociada a la aplicación lineal:

$$\begin{aligned} f(\vec{u}_1) &= f(2, 0, 0) = (2, 0, 0) = 2 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \\ f(\vec{u}_2) &= f(-1, 3, 0) = (2, 3, 0) = 2 \cdot \vec{e}_1 + 3 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \\ f(\vec{u}_3) &= f(0, 0, 5) = (0, 0, 0) = 0 \cdot \vec{e}_1 + 0 \cdot \vec{e}_2 + 0 \cdot \vec{e}_3 \end{aligned}$$

es decir,

$$M(f)_{B,C} = \begin{pmatrix} 2 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Finalmente si en ambos espacios consideremos la base $B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, con $\vec{u}_1, \vec{u}_2, \vec{u}_3$ dados arriba, tenemos

$$\begin{aligned} f(\vec{u}_1) &= f(2, 0, 0) = (2, 0, 0) = \alpha_{11}(2, 0, 0) + \alpha_{12}(-1, 3, 0) + \alpha_{13}(0, 0, 5) \\ f(\vec{u}_2) &= f(-1, 3, 0) = (2, 3, 0) = \alpha_{21}(2, 0, 0) + \alpha_{22}(-1, 3, 0) + \alpha_{23}(0, 0, 5) \\ f(\vec{u}_3) &= f(0, 0, 5) = (0, 0, 0) = \alpha_{31}(2, 0, 0) + \alpha_{32}(-1, 3, 0) + \alpha_{33}(0, 0, 5) \end{aligned}$$

es decir,

$$\begin{cases} 2 = 2\alpha_{11} - \alpha_{12} \\ 0 = 3\alpha_{12} \\ 0 = 5\alpha_{13} \end{cases} \implies \alpha_{11} = 1, \quad \alpha_{12} = 0, \quad \alpha_{13} = 0$$

$$\begin{cases} 2 = 2\alpha_{21} - \alpha_{22} \\ 3 = 3\alpha_{22} \\ 0 = 5\alpha_{23} \end{cases} \implies \alpha_{21} = \frac{3}{2}, \quad \alpha_{22} = 1, \quad \alpha_{23} = 0$$

$$\begin{cases} 0 = 2\alpha_{31} - \alpha_{32} \\ 0 = 3\alpha_{32} \\ 0 = 5\alpha_{33} \end{cases} \implies \alpha_{31} = 0, \quad \alpha_{32} = 0, \quad \alpha_{33} = 0$$

8.3. MATRIZ ASOCIADA A UNA APLICACIÓN LINEAL.

y entonces

$$M(f)_{B,B} = \begin{pmatrix} 1 & 3/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Está claro con este ejemplo que la matriz asociada a una aplicación lineal depende de las bases que tengamos en el espacio inicial y en el final. Pero una vez fijadas las bases la matriz es única.

Pasamos ahora a estudiar la matriz de cambio de base.

Definición. Sean $A = \{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}$ y $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ dos bases de un espacio vectorial $(E, +, \cdot)$ de dimensión n sobre un cuerpo K . Los vectores de la base B se pueden expresar en función de la base A

$$\begin{cases} \vec{b}_1 = a_{11}\vec{a}_1 + a_{21}\vec{a}_2 + \dots + a_{n1}\vec{a}_n \\ \vec{b}_2 = a_{12}\vec{a}_1 + a_{22}\vec{a}_2 + \dots + a_{n2}\vec{a}_n \\ \vdots \\ \vec{b}_n = a_{1n}\vec{a}_1 + a_{2n}\vec{a}_2 + \dots + a_{nn}\vec{a}_n. \end{cases}$$

La matriz regular (esto es $\det P \neq 0$), de orden n

$$P = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (8.3)$$

se llama **matriz de cambio de base** y tiene por columnas las coordenadas de la nueva base B en función de la vieja base A .

Sea ahora $\vec{x} \in E$ un vector cualquiera; se podrá escribir este vector en función de las dos bases como:

$$\vec{x} = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \dots + \alpha_n \vec{a}_n \quad \vec{x} = \beta_1 \vec{b}_1 + \beta_2 \vec{b}_2 + \dots + \beta_n \vec{b}_n.$$

Sustituyendo las expresiones de los vectores de la nueva base en función de los viejos, tendremos

$$\begin{aligned} \vec{x} &= \beta_1 \vec{b}_1 + \beta_2 \vec{b}_2 + \dots + \beta_n \vec{b}_n = \beta_1 (a_{11}\vec{a}_1 + a_{21}\vec{a}_2 + \dots + a_{n1}\vec{a}_n) + \\ &\quad \beta_2 (a_{12}\vec{a}_1 + a_{22}\vec{a}_2 + \dots + a_{n2}\vec{a}_n) + \dots + \beta_n (a_{1n}\vec{a}_1 + a_{2n}\vec{a}_2 + \dots + a_{nn}\vec{a}_n) = \\ &= (\beta_1 a_{11} + \beta_2 a_{12} + \dots + \beta_n a_{1n}) \vec{a}_1 + (\beta_1 a_{21} + \beta_2 a_{22} + \dots + \beta_n a_{2n}) \vec{a}_2 + \dots + \\ &\quad (\beta_1 a_{n1} + \beta_2 a_{n2} + \dots + \beta_n a_{nn}) \vec{a}_n \end{aligned}$$

y como la expresión de un vector en una determinada base es única, tendremos

$$\begin{cases} \beta_1 a_{11} + \beta_2 a_{12} + \dots + \beta_n a_{1n} = \alpha_1 \\ \beta_1 a_{21} + \beta_2 a_{22} + \dots + \beta_n a_{2n} = \alpha_2 \\ \vdots \\ \beta_1 a_{n1} + \beta_2 a_{n2} + \dots + \beta_n a_{nn} = \alpha_n \end{cases} \Leftrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix},$$

es decir

$$P\bar{\beta} = \bar{\alpha}$$

donde P es la matriz de cambio de base. $\bar{\beta}$ la matriz columna formada por las componentes de \vec{x} en la nueva base, y $\bar{\alpha}$ la matriz columna formada por las componentes de \vec{x} en la vieja base.

Estudiemos, finalmente, como se transforman las matrices asociadas en distintas bases a una misma aplicación lineal.

Sean $(E, +, \cdot)$, $\dim E = n$, y $(F, +, \cdot)$, $\dim F = m$, dos espacios vectoriales sobre un cuerpo K , y sean $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ y $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ dos bases de E y F , respectivamente. Sea $f: E \rightarrow F$ una aplicación lineal entre dichos espacios vectoriales tal que, para todo $\vec{x} \in E$ existe $\vec{y} = f(\vec{x}) \in F$.

Sabemos que respecto a dichas bases, hay una y sólo una matriz asociada a dicha aplicación, la matriz (8.2),

$$M(f)_{B_1, B_2} = A \quad \text{tal que} \quad A\bar{x} = \bar{y},$$

donde \bar{x} representa la matriz columna formada por las componentes del vector \vec{x} en la base B_1 , y así mismo \bar{y} representa la matriz columna formada por las componentes del vector \vec{y} en la base B_2 .

De igual forma, si consideramos en E y F dos nuevas bases $B'_1 = \{\vec{u}'_1, \vec{u}'_2, \dots, \vec{u}'_n\}$ y $B'_2 = \{\vec{v}'_1, \vec{v}'_2, \dots, \vec{v}'_m\}$, respecto a estas nuevas bases, hay una y sólo una matriz asociada

$$M(f)_{B'_1, B'_2} = B \quad \text{tal que} \quad B\bar{x}' = \bar{y}'.$$

Como en E hemos realizado un cambio de base, existe una matriz cambio de base P , ver (8.3), regular, $n \times n$, tal que $P\bar{x}' = \bar{x}$. De la misma manera, como en F hemos también realizado un cambio de base, existe una matriz cambio de base Q , regular, $m \times m$, tal que $Q\bar{y}' = \bar{y}$. Luego

$$\begin{cases} \bar{y} = A\bar{x} \\ \bar{y} = Q\bar{y}' \\ \bar{x} = P\bar{x}' \end{cases} \Rightarrow Q\bar{y}' = A\bar{x} = AP\bar{x}' \Rightarrow Q^{-1}Q\bar{y}' = \bar{y}' = Q^{-1}AP\bar{x}',$$

esto implica por la unicidad de la matriz asociada a una aplicación lineal

$$B = Q^{-1}AP \quad \text{con un esquema} \quad \begin{array}{ccc} f: E & \rightarrow & F \\ B_1 & \xrightarrow{A} & B_2 \\ P \downarrow & & \downarrow Q \\ B'_1 & \xrightarrow{B} & B'_2 \end{array}$$

En el caso $E = F$, será $P = Q$, y por lo tanto $B = P^{-1}AP$.

Ejemplo. Sea la aplicación lineal $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $f(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3)$. Halla la matriz de f respecto a las bases

$$B'_1 = \{(1, 1, 0), (0, 1, -1), (1, 0, 0)\} \quad \text{y} \quad B'_2 = \{(1, 2), (0, 1)\}.$$

8.3. MATRIZ ASOCIADA A UNA APLICACIÓN LINEAL.

Solución:

$$B = Q^{-1}AP \quad \text{esquema} \quad \begin{array}{ccc} f: \mathbb{R}^3 \rightarrow \mathbb{R}^2 & & \\ \text{base canónica} \xrightarrow{A} \text{base canónica} & & \\ P \downarrow & & \downarrow Q \\ B'_1 & \xrightarrow{B} & B'_2 \end{array}$$

$$\begin{cases} f(1, 0, 0) = (1, 0) \\ f(0, 1, 0) = (1, 1) \\ f(0, 0, 1) = (0, 1) \end{cases} \Rightarrow A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$\begin{cases} (1, 1, 1) = 1 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) + 1 \cdot (0, 0, 1) \\ (0, 1, -1) = 0 \cdot (1, 0, 0) + 1 \cdot (0, 1, 0) - 1 \cdot (0, 0, 1) \\ (1, 0, 0) = 1 \cdot (1, 0, 0) + 0 \cdot (0, 1, 0) + 0 \cdot (0, 0, 1) \end{cases} \Rightarrow P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

$$\begin{cases} (1, 2) = 1 \cdot (1, 0) + 2 \cdot (0, 1) \\ (0, 1) = 0 \cdot (1, 0) + 1 \cdot (0, 1) \end{cases} \Rightarrow Q = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \Rightarrow Q^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}.$$

Entonces

$$\begin{aligned} B &= Q^{-1}AP = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 1 & 0 \\ -2 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ -2 & -2 & -2 \end{pmatrix}. \end{aligned}$$

Ejercicios.

- 1) Dada la aplicación $f : (\mathbb{R}^3, +, \cdot) \rightarrow (\mathbb{R}^4, +, \cdot)$, $f(x_1, x_2, x_3) = (x_1, x_2, x_3 - x_1, x_2)$, halla la matriz de dicha aplicación.
- 2) Dado el homomorfismo $f : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ definido por $f(1, 0, 1) = (1, 0, 1, 0)$, $f(0, 1, 1) = (0, 1, 1, 0)$, $f(1, 1, 0) = (1, 1, 1, 1)$. Halla la matriz de dicho homomorfismo.
- 3) Dada la aplicación $f : (\mathbb{R}^2, +, \cdot) \rightarrow (\mathbb{R}^4, +, \cdot)$ de la cual se sabe que las imágenes de los vectores $(2, -1)$ y $(4, 1)$ son respectivamente $(1, 0, -1, 3)$ y $(2, -2, 3, 1)$. Halla la matriz que define la aplicación, si todos los vectores están referidos en ambos espacios a las bases canónicas.
- 4) Sea f una aplicación de \mathbb{R}^3 en si mismo, y B una base de \mathbb{R}^3 formada por los vectores $\vec{u}_1, \vec{u}_2, \vec{u}_3$, tales que $f(\vec{u}_1) = \vec{u}_1 + 2\vec{u}_2 - \vec{u}_3$, $f(\vec{u}_2) = -3\vec{u}_1 - \vec{u}_2 + 5\vec{u}_3$, $f(\vec{u}_3) = -2\vec{u}_1 + \vec{u}_2 + 2\vec{u}_3$. Halla la matriz asociada a la aplicación f respecto de la base $B_1 = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ donde $\vec{v}_1 = \vec{u}_2 + \vec{u}_3$, $\vec{v}_2 = \vec{u}_1 + \vec{u}_3$ y $\vec{v}_3 = \vec{u}_1 + \vec{u}_2$.

8.4 Suma, producto y composición de aplicaciones lineales.

Definición. Sean E y F dos espacios vectoriales sobre el cuerpo K . Al conjunto de las aplicaciones lineales de E en F se llama $\mathcal{L}_K(E, F)$, o bien, si no es necesario indicar el cuerpo, se representa simplemente por $\mathcal{L}(E, F)$. En particular, si $E = F$, tenemos el conjunto de los endomorfismos en E , que se representa por $\mathcal{L}(E)$.

Definición. Para todas $f, g \in \mathcal{L}(E, F)$ se define la aplicación **suma de las aplicaciones lineales** f y g : $(f + g) : E \rightarrow F$, de manera que $(f + g)(\vec{x}) = f(\vec{x}) + g(\vec{x})$ para todo vector $\vec{x} \in E$.

Veamos que $(f + g)$ es también una aplicación lineal de E en F . $\forall \alpha, \beta \in K, \forall \vec{x}, \vec{y} \in E$.

$$\begin{aligned} (f + g)(\alpha \vec{x} + \beta \vec{y}) &= f(\alpha \vec{x} + \beta \vec{y}) + g(\alpha \vec{x} + \beta \vec{y}) = && \text{por definición} \\ &= f(\alpha \vec{x}) + f(\beta \vec{y}) + g(\alpha \vec{x}) + g(\beta \vec{y}) = && f \text{ y } g \text{ lineales} \\ &= \alpha [f(\vec{x}) + g(\vec{x})] + \beta [f(\vec{y}) + g(\vec{y})] = \\ &= \alpha [(f + g)(\vec{x})] + \beta [(f + g)(\vec{y})], && \text{por definición} \end{aligned}$$

luego $(f + g)$ es efectivamente también una aplicación lineal de E en F .

Se puede demostrar que $(\mathcal{L}(E, F), +)$ tiene una estructura de grupo conmutativo.

También se puede demostrar que si $M(f)_{B_1 B_2}$ y $M(g)_{B_1 B_2}$ son las matrices asociadas a las aplicaciones f y g en las bases B_1 y B_2 , de E y F , respectivamente, entonces la matriz asociada a la aplicación lineal suma $(f + g)$ en las bases B_1 y B_2 , es la suma de las matrices $M(f)_{B_1 B_2}$ y $M(g)_{B_1 B_2}$: $M(f)_{B_1 B_2} + M(g)_{B_1 B_2}$.

Definición. Para todo escalar $\lambda \in K$ y toda aplicación lineal $f \in \mathcal{L}(E, F)$ se define la aplicación **producto de una aplicación lineal f por un escalar λ** : $(\lambda \cdot f) : E \rightarrow F$, de manera que $(\lambda \cdot f)(\vec{x}) = \lambda [f(\vec{x})]$ para todo vector $\vec{x} \in E$.

Veamos que $(\lambda \cdot f)$ es también una aplicación lineal de E en F . $\forall \alpha, \beta \in K, \forall \vec{x}, \vec{y} \in E$.

$$\begin{aligned} (\lambda \cdot f)(\alpha \vec{x} + \beta \vec{y}) &= \lambda f(\alpha \vec{x} + \beta \vec{y}) = && \text{por definición} \\ &= \lambda f(\alpha \vec{x}) + \lambda f(\beta \vec{y}) = && f \text{ lineal} \\ &= \alpha \lambda [f(\vec{x})] + \beta \lambda [f(\vec{y})] = \\ &= \alpha [(\lambda \cdot f)(\vec{x})] + \beta [(\lambda \cdot f)(\vec{y})], && \text{por definición} \end{aligned}$$

luego $(\lambda \cdot f)$ es efectivamente también una aplicación lineal de E en F .

Se puede demostrar (pero no lo haremos) que $(\mathcal{L}(E, F), +, \cdot)$ tiene una estructura de espacio vectorial.

También se puede demostrar que si $M(f)_{B_1 B_2}$ es la matriz asociada a la aplicación f en las bases B_1 y B_2 , de E y F , respectivamente, luego la matriz asociada a la aplicación lineal producto por un escalar $(\lambda \cdot f)$ en las bases B_1 y B_2 , es el producto de la matriz $M(f)_{B_1 B_2}$ por el escalar λ : $\lambda \cdot M(f)_{B_1 B_2}$.

8.5. DETERMINANTE DE UNA MATRIZ $N \times N$. PROPIEDADES.

Sean E, F y H espacios vectoriales sobre el cuerpo K , de dimensión p, n y q , respectivamente. Sean $f : E \rightarrow F$ una aplicación lineal de E en F , $g : F \rightarrow H$ una aplicación lineal de F en H .

Definición. Para todas aplicaciones lineales $f \in \mathcal{L}(E, F)$ y $g \in \mathcal{L}(F, H)$ se define la aplicación **composición de las aplicaciones lineales** f y g : $(g \circ f) : E \rightarrow H$, de manera que $(g \circ f)(\vec{x}) = g(f(\vec{x}))$ para todo vector $\vec{x} \in E$.

Veamos que $(g \circ f)$ es también una aplicación lineal de E en H . $\forall \alpha, \beta \in K, \forall \vec{x}, \vec{y} \in E$,

$$\begin{aligned} (g \circ f)(\alpha \vec{x} + \beta \vec{y}) &= g(f(\alpha \vec{x} + \beta \vec{y})) = && \text{por definición} \\ &= g(f(\alpha \vec{x}) + f(\beta \vec{y})) = && f \text{ lineal} \\ &= g(\alpha f(\vec{x}) + \beta f(\vec{y})) = && f \text{ lineal} \\ &= \alpha g(f(\vec{x})) + \beta g(f(\vec{y})) = && g \text{ lineal} \\ &= \alpha (g \circ f)(\vec{x}) + \beta (g \circ f)(\vec{y}), && \text{por definición} \end{aligned}$$

luego $(g \circ f)$ es una aplicación lineal de E en H .

Tendremos que si $M(f)_{B_1, B_2}$ es la matriz asociada a la aplicación f en las bases B_1 y B_2 , de E y F , respectivamente, y $M(g)_{B_2, B_3}$ es la matriz asociada a la aplicación g en las bases B_2 y B_3 , de F y H , respectivamente, entonces la matriz asociada a la aplicación lineal composición de las aplicaciones lineales f y g en las bases B_1 y B_3 , es el producto de las matrices $M(f)_{B_1, B_2}$ por $M(g)_{B_2, B_3}$: $M(g)_{B_2, B_3} \cdot M(f)_{B_1, B_2}$.

Ejercicios.

- 1) Dadas las dos aplicaciones lineales $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definidas por $f(x, y) = (y, x)$ y $g(x, y) = (3x, 3x + y)$, calcula la suma $f + g$ y los productos $3f(x, y)$ y $9g(x, y)$.
- 2) Dadas las dos aplicaciones lineales $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definidas por $f(x, y) = (y, x)$ y $g(x, y) = (3x, 3x + y)$, calcula las composiciones $g \circ f$ y $f \circ g$.

8.5 Determinante de una matriz $n \times n$. Propiedades.

Un concepto fundamental del álgebra lineal es el **determinante** de una matriz $n \times n$. El único enfoque que se dará aquí para calcular el determinante será la definición recursiva. El determinante de una matriz A se denotará por

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Definición. Una **submatriz** de una matriz A es una matriz *extraída* de A suprimiendo algunas filas y/o columnas de A .

Definición. Si $A = (a)$ es una matriz 1×1 , entonces $\det A = |A| = a$.

Definición. El **menor complementario**, ó sólo **menor**, M_{ij} es el determinante de la submatriz $(n-1) \times (n-1)$ de una matriz A , $n \times n$, obtenida suprimiendo la i -ésima fila y la j -ésima columna.

Definición. El **cofactor** (ó también **adjunto** del elemento a_{ij}), A_{ij} asociado con M_{ij} se define como $A_{ij} = (-1)^{i+j} M_{ij}$.

Definición. El **determinante** de una matriz A , $n \times n$, donde $n > 1$, está dado ya sea por

$$\det A = \sum_{j=1}^n a_{ij} A_{ij} \quad \text{para cualquier } i = 1, 2, \dots, n,$$

o por

$$\det A = \sum_{i=1}^n a_{ij} A_{ij} \quad \text{para cualquier } j = 1, 2, \dots, n.$$

Es decir, el valor de un determinante es igual a la suma de los productos de los elementos de una línea o de una columna por sus cofactores correspondientes.

Veamos un ejemplo con una matriz que tiene dos ceros en la cuarta columna y dos en la tercera fila; por comparación veremos cómo resulta más comodo desarrollar por las líneas que más ceros tienen. Si desarrollamos por la primera fila tenemos:

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & -3 & 1 \\ 6 & 0 & 4 & 0 \\ -1 & 3 & 2 & 0 \end{vmatrix} &= 1 \cdot (-1)^{1+1} \begin{vmatrix} 5 & -3 & 1 \\ 0 & 4 & 0 \\ 3 & 2 & 0 \end{vmatrix} + 2 \cdot (-1)^{1+2} \begin{vmatrix} 0 & -3 & 1 \\ 6 & 4 & 0 \\ -1 & 2 & 0 \end{vmatrix} + \\ &+ 3 \cdot (-1)^{1+3} \begin{vmatrix} 0 & 5 & 1 \\ 6 & 0 & 0 \\ -1 & 3 & 0 \end{vmatrix} + 4 \cdot (-1)^{1+4} \begin{vmatrix} 0 & 5 & -3 \\ 6 & 0 & 4 \\ -1 & 3 & 2 \end{vmatrix} = \\ &= 1 \cdot (-12) - 2 \cdot (12 + 4) + 3 \cdot (18) - 4 \cdot (-20 - 54 - 60) = \\ &= -12 - 32 + 54 + 536 = 546; \end{aligned}$$

y si desarrollamos por la cuarta columna obtenemos

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & -3 & 1 \\ 6 & 0 & 4 & 0 \\ -1 & 3 & 2 & 0 \end{vmatrix} &= 4 \cdot (-1)^{1+4} \begin{vmatrix} 0 & 5 & -3 \\ 6 & 0 & 4 \\ -1 & 3 & 2 \end{vmatrix} + 1 \cdot (-1)^{2+4} \begin{vmatrix} 1 & 2 & 3 \\ 6 & 0 & 4 \\ -1 & 3 & 2 \end{vmatrix} + 0 + 0 = \\ &= -4 \cdot (-20 - 54 - 60) + (-8 + 54 - 12 - 24) = 546. \end{aligned}$$

Usando inducción matemática, se puede demostrar que, si $n > 1$, el uso de las definiciones dadas para calcular el determinante de una matriz, en general $n \times n$, requiere $n!$ multiplicaciones / divisiones y de $(n-1)$ sumas / restas. Incluso para valores relativamente pequeños de n , el número de cálculos puede llegar a ser inmanejable. Para matrices 2×2 y 3×3 las cosas son relativamente sencillas.

8.5. DETERMINANTE DE UNA MATRIZ $N \times N$. PROPIEDADES.

- En el caso de una matriz de orden 2 tenemos que el determinante es igual al producto de los elementos de la diagonal principal menos el producto de los elementos de la diagonal secundaria:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}.$$

Por ejemplo, para las matrices $A = \begin{pmatrix} 2 & 7 \\ -1 & -5 \end{pmatrix}$ y $B = \begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$,

$$\det A = \begin{vmatrix} 2 & 7 \\ -1 & -5 \end{vmatrix} = 2 \cdot (-5) - (-1) \cdot 7 = -10 + 7 = -3;$$

$$\det B = \begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix} = 4 \cdot 3 - 2 \cdot 1 = 12 - 2 = 10.$$

- En el caso de matriz de orden 3 tenemos la **regla de Sarrus**: los términos positivos están formados por el producto de los elementos de la diagonal principal y por los productos de los elementos que constituyen una línea paralela a ella multiplicados por el vértice opuesto; los términos negativos están formados por el producto de los elementos de la diagonal secundaria y por los productos de los elementos que constituyen la paralela a ella multiplicados por los elementos del vértice opuesto.

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} + \\ - a_{31} \cdot a_{22} \cdot a_{13} - a_{32} \cdot a_{23} \cdot a_{11} - a_{33} \cdot a_{21} \cdot a_{12}.$$

Por ejemplo, dada la matriz $A = \begin{pmatrix} 1 & 3 & -2 \\ 0 & 3 & 2 \\ -2 & 1 & 4 \end{pmatrix}$,

$$\det A = \begin{vmatrix} 1 & 3 & -2 \\ 0 & 3 & 2 \\ -2 & 1 & 4 \end{vmatrix} = \\ = 1 \cdot 3 \cdot 4 + 3 \cdot 2 \cdot (-2) + 0 \cdot 1 \cdot (-2) - (-2) \cdot 3 \cdot (-2) - 1 \cdot 2 \cdot 1 - 0 \cdot 3 \cdot 4 \\ = 12 - 12 + 0 - 12 - 2 - 0 = -14.$$

Pasamos ahora a enumerar las **propiedades generales** de los determinantes.
Sea A una matriz $n \times n$:

1. Si cualquier fila o columna de A tiene sólo componentes cero, entonces $\det A = 0$; por ejemplo

$$\det A = \begin{vmatrix} 0 & 0 \\ -1 & -5 \end{vmatrix} = 0 \cdot (-5) - (-1) \cdot 0 = 0;$$

en este otro ejemplo, desarrollamos por la fila que tiene ceros

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & -3 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 3 & 2 & 0 \end{vmatrix} = 0 \cdot (-1)^{3+1} \begin{vmatrix} 2 & 3 & 4 \\ 5 & -3 & 1 \\ 3 & 2 & 0 \end{vmatrix} + 0 \cdot (-1)^{3+2} \begin{vmatrix} 1 & 3 & 4 \\ 0 & -3 & 1 \\ -1 & 2 & 0 \end{vmatrix} + \\ + 0 \cdot (-1)^{3+3} \begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 1 \\ -1 & 3 & 0 \end{vmatrix} + 0 \cdot (-1)^{3+4} \begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -3 \\ -1 & 3 & 2 \end{vmatrix} = 0.$$

El mismo resultado se obtiene desarrollando por cualquier otra fila o columna.

2. El valor de un determinante no varía cuando se cambian ordenadamente sus filas por sus columnas o viceversa, esto es $|A| = |A^t|$; por ejemplo

$$\det A = \begin{vmatrix} 2 & 3 & 1 \\ 4 & -4 & 0 \\ 2 & 1 & -1 \end{vmatrix} = 8 + 4 + 4 - (-8) - 0 - (-12) = 32;$$

$$\det A^t = \begin{vmatrix} 2 & 4 & 2 \\ 3 & -4 & 1 \\ 1 & 0 & -1 \end{vmatrix} = 8 + 0 + 4 - (-8) - (-12) - 0 = 32;$$

3. Si en un determinante se cambian entre sí dos líneas paralelas (filas o columnas) el valor absoluto del determinante no varía pero cambia de signo; por ejemplo, considerando la matriz A del ejemplo anterior y denotando con \tilde{A} la nueva matriz en la cual se han cambiado entre sí la primera y la segunda fila, tendremos

$$\det \tilde{A} = \begin{vmatrix} 4 & -4 & 0 \\ 2 & 3 & 1 \\ 2 & 1 & -1 \end{vmatrix} = -12 - 8 + 0 - 0 - 4 - 8 = -32;$$

4. Si un determinante tiene dos líneas paralelas (filas o columnas) iguales, entonces su valor es nulo; por ejemplo

$$\det A = \begin{vmatrix} 4 & -4 & 4 \\ 2 & 3 & 2 \\ 2 & 1 & 2 \end{vmatrix} = 24 - 16 + 8 - 24 - 8 + 16 = 0;$$

5. Si en un determinante se multiplican los elementos de una línea (fila o columna) por un escalar α , el determinante queda multiplicado por el mismo valor α ; por ejemplo

$$\det A = \begin{vmatrix} 3 \cdot 2 & 3 & 1 \\ 3 \cdot 4 & -4 & 0 \\ 3 \cdot 2 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 6 & 3 & 1 \\ 12 & -4 & 0 \\ 6 & 1 & -1 \end{vmatrix} = \\ = 24 + 0 + 12 - (-24) - 0 - (-36) = 96 = 3 \cdot 32;$$

8.5. DETERMINANTE DE UNA MATRIZ $N \times N$. PROPIEDADES.

claramente si queremos multiplicar un número por un determinante basta multiplicar ese número por los elementos de una línea:

$$3 \cdot \begin{vmatrix} 2 & 3 & 1 \\ 4 & -4 & 0 \\ 2 & 1 & -1 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 1 \\ 3 \cdot 4 & 3 \cdot (-4) & 3 \cdot 0 \\ 2 & 1 & -1 \end{vmatrix} = 96 = 3 \cdot 32;$$

6. Si un determinante tiene dos líneas paralelas (filas o columnas) proporcionales, su valor es nulo; por ejemplo

$$\begin{vmatrix} 4 & -4 & 8 \\ 2 & 3 & 4 \\ 2 & 1 & 4 \end{vmatrix} = \begin{vmatrix} 4 & -4 & 2 \cdot 4 \\ 2 & 3 & 2 \cdot 2 \\ 2 & 1 & 2 \cdot 2 \end{vmatrix} = 2 \cdot \begin{vmatrix} 4 & -4 & 4 \\ 2 & 3 & 2 \\ 2 & 1 & 2 \end{vmatrix} = 2 \cdot 0 = 0;$$

7. La suma de los productos de los elementos de una línea por los cofactores de una línea paralela a ella vale cero; por ejemplo, en la siguiente matriz hacemos primera columna por los cofactores de la tercera columna:

$$\begin{vmatrix} 1 & 2 & 3 \\ 6 & 0 & 4 \\ -1 & 3 & 2 \end{vmatrix} = 1 \cdot (-1)^{1+3} \cdot \begin{vmatrix} 6 & 0 \\ -1 & 3 \end{vmatrix} + 6 \cdot (-1)^{2+3} \cdot \begin{vmatrix} 1 & 2 \\ -1 & 3 \end{vmatrix} + \\ + (-1) \cdot (-1)^{3+3} \cdot \begin{vmatrix} 1 & 2 \\ 6 & 0 \end{vmatrix} = 1 \cdot 18 - 6 \cdot (3 + 2) - 1 \cdot (-12) = \\ = 18 - 30 + 12 = 0;$$

8. Si cada uno de los elementos de una línea está formado por la suma de h sumandos, entonces dicho determinante se puede descomponer en la suma de h determinantes que tienen las restantes líneas iguales y en lugar de aquella, otra formada por los primeros, segundos, ..., h sumandos. Por ejemplo,

$$\begin{vmatrix} 3+2+5 & -1+7 & 5+2+4 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} = \begin{vmatrix} 3 & -1 & 5 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 7 & 2 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} + \begin{vmatrix} 5 & 0 & 4 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix}$$

es decir, por un lado

$$\begin{vmatrix} 10 & 6 & 11 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} = -50 + 84 + 132 + 77 - 80 - 90 = 73$$

y por el otro

$$\begin{vmatrix} 3 & -1 & 5 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 7 & 2 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} + \begin{vmatrix} 5 & 0 & 4 \\ 3 & -1 & 2 \\ 7 & 4 & 5 \end{vmatrix} = \\ = [-15 - 14 + 60 + 35 - 24 + 15] + [-10 + 98 + 24 + 14 - 16 - 105] + \\ + [-25 + 0 + 48 + 28 - 40 - 0] = 57 + 5 + 11 = 73;$$

9. Si a un determinante, en una de sus líneas, se le añade una combinación lineal de otras paralelas a ella, el determinante no varía; por ejemplo, en el siguiente determinante a la primera fila se le añade dos veces la segunda fila y se le resta la tercera fila; luego se le aplica la propiedad 8:

$$\begin{aligned} \begin{vmatrix} 2 & 7 & -5 \\ -1 & 3 & 0 \\ 4 & -2 & 10 \end{vmatrix} &= \begin{vmatrix} 2+2 \cdot (-1) - 4 & 7+2 \cdot 3 + 2 & -5+2 \cdot 0 - 10 \\ -1 & 3 & 0 \\ 4 & -2 & 10 \end{vmatrix} = \\ &= \begin{vmatrix} 2 & 7 & -5 \\ -1 & 3 & 0 \\ 4 & -2 & 10 \end{vmatrix} + \begin{vmatrix} -2 & 6 & 0 \\ -1 & 3 & 0 \\ 4 & -2 & 10 \end{vmatrix} + \begin{vmatrix} -4 & 72 & -10 \\ -1 & 3 & 0 \\ 4 & -2 & 10 \end{vmatrix} \end{aligned}$$

finalmente los últimos dos determinantes son nulos por la propiedad 6.

10. Si en un determinante una de sus líneas es combinación lineal de otras paralelas a ella, el determinante es nulo; por ejemplo en el determinante que sigue, la tercera columna es igual a dos veces la primera columna más la segunda columna

$$\begin{aligned} \begin{vmatrix} 1 & -4 & -2 \\ 2 & 7 & 11 \\ 3 & 6 & 12 \end{vmatrix} &= \begin{vmatrix} 1 & -4 & 2 \cdot 1 - 4 \\ 2 & 7 & 2 \cdot 2 + 7 \\ 3 & 6 & 2 \cdot 3 + 6 \end{vmatrix} = \\ &= \begin{vmatrix} 1 & -4 & 2 \cdot 1 \\ 2 & 7 & 2 \cdot 2 \\ 3 & 6 & 2 \cdot 3 \end{vmatrix} + \begin{vmatrix} 1 & -4 & -4 \\ 2 & 7 & 7 \\ 3 & 6 & 6 \end{vmatrix} = 0 + 0; \end{aligned}$$

estos últimos determinantes son nulos por las propiedades 6 y 4, respectivamente.

11. Si $A, B \in M_{n \times n}$ luego

$$\det(A \cdot B) = \det A \cdot \det B.$$

Por ejemplo,

$$\det A = \begin{vmatrix} 2 & 7 \\ -1 & -5 \end{vmatrix} = -3, \quad \det B = \begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix} = 10, \quad \det(A \cdot B) = \begin{vmatrix} 15 & 25 \\ -9 & -17 \end{vmatrix} = -30.$$

Ejercicio.

Demuestra sin desarrollar que

a) $\begin{vmatrix} 1 & a & b+c \\ 1 & b & c+a \\ 1 & c & a+b \end{vmatrix} = 0,$

b) $\begin{vmatrix} 3 & 15 & 1 \\ 5 & 25 & 0 \\ 1 & 5 & 7 \end{vmatrix} = 0,$

c) $\begin{vmatrix} a^2 & ab & b^2 \\ 2a & a+b & 2b \\ 1 & 1 & 1 \end{vmatrix} = (a-b)^3,$

d) $\begin{vmatrix} 3 & 5 & 6 \\ 7 & 5 & 4 \\ 10 & 10 & 10 \end{vmatrix} = 0.$

8.6 Cálculo de determinantes.

Veamos ahora algunos métodos para desarrollar determinantes de cualquier orden.

1. Por **adjuntos** o cofactores. Este es el método que hemos ya estudiado: el valor del determinante es igual a la suma de los productos de los elementos de una línea o de una columna por sus cofactores correspondientes. Además de los ejemplos ya vistos, otro ejemplo es el siguiente, en el cual desarrollamos por cofactores de la cuarta fila.

$$\begin{aligned} \begin{vmatrix} 1 & 4 & 5 & 1 \\ 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} &= 1 \cdot (-1)^{4+1} \begin{vmatrix} 4 & 5 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{vmatrix} + 1 \cdot (-1)^{4+3} \begin{vmatrix} 1 & 4 & 1 \\ 3 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix} + \\ &+ 1 \cdot (-1)^{4+4} \begin{vmatrix} 1 & 4 & 5 \\ 3 & 1 & 1 \\ 1 & 2 & 1 \end{vmatrix} = -(4 + 1 - 2 - 5) + \\ &-(1 + 6 - 1 - 12) + (1 + 4 + 30 - 5 - 2 - 12) = 24. \end{aligned}$$

2. **Regla de Chio.** Se busca una línea que tenga algún elemento con valor uno, que se denomina **pivote**. Mediante combinaciones lineales de la fila o columna en que está situado el pivote, se crean ceros en las restantes posiciones de la misma columna o fila, y luego se desarrolla el determinante por los cofactores de esa columna o fila. Al ser todos los elementos de esa línea ceros salvo el pivote, el determinante queda reducido al cálculo de un solo determinante de menor orden que el inicial.

Si no existe ningún elemento que valga uno, se elige un elemento cualquiera, suponemos que sea el $K = a_{ij}$, entonces se pueden dividir todos los elementos de la fila i o columna j por el valor K , convirtiendo el término a_{ij} en uno. No tenemos que olvidar que procediendo así el determinante quedará multiplicado externamente por el valor K .

Hagamos por este método el determinante anterior, eligiendo como elemento pivote el $a_{44} = 1$ y creando ceros en la cuarta columna con los siguientes pasos: se sustituyen las filas con las operaciones $(E_1 - E_4) \rightarrow (E_1)$ y $(E_3 - E_4) \rightarrow (E_3)$. [Nótese que $(E_i + \lambda E_j) \rightarrow (E_i)$ quiere decir que la fila E_j se multiplica por la constante λ , se la suma a la fila E_i , y se usa la fila resultante en lugar de E_i]. Luego se desarrolla el determinante por la cuarta columna (podría también hacerse por la tercera fila) y finalmente se desarrolla el determinante de orden dos por la tercera fila.

$$\begin{aligned} \begin{vmatrix} 1 & 4 & 5 & 1 \\ 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} &= \begin{vmatrix} 0 & 2 & 4 & 0 \\ 3 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{vmatrix} = 1 \cdot (-1)^{4+4} \begin{vmatrix} 0 & 2 & 4 \\ 3 & 1 & 1 \\ 0 & 2 & 0 \end{vmatrix} = \\ &= 2 \cdot (-1)^{3+2} \begin{vmatrix} 0 & 4 \\ 3 & 1 \end{vmatrix} = -2 \cdot (-12) = 24. \end{aligned}$$

3. **Triangulando la matriz** (se suele usar para programar con ordenador). Al paso i cualquiera, se deja fija la fila i -ésima, y tomando como pivote el elemento a_{ii} , se van haciendo cero todos los elementos de su columna que estén por debajo de él. Si ocurre que el elemento a_{ii} es cero, se intercambia primero su fila con alguna otra inferior, *sin olvidarse de cambiar el signo del determinante*, y si esto no es posible por ser todos los elementos de su columna por debajo del a_{ii} nulos, incluido este, el determinante vale cero por tener una columna de ceros.

Volvemos a considerar el ejemplo anterior. Para crear ceros en la primera columna por debajo del elemento a_{11} realizamos las siguientes operaciones:

$$(E_2 - 3E_1) \rightarrow (E_2), \quad (E_3 - E_1) \rightarrow (E_3) \quad \text{y} \quad (E_4 - E_1) \rightarrow (E_4);$$

para anular los elementos de la segunda columna por debajo del elemento a_{22} haremos

$$(E_3 - \frac{2}{11}E_2) \rightarrow (E_3) \quad \text{y} \quad (E_4 - \frac{4}{11}E_2) \rightarrow (E_4);$$

y finalmente para el único elemento que nos queda en la tercera columna por debajo del elemento a_{33} haremos

$$(E_4 + \frac{12}{16}E_3) \rightarrow (E_4).$$

$$\begin{aligned} \begin{vmatrix} 1 & 4 & 5 & 1 \\ 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{vmatrix} &= \begin{vmatrix} 1 & 4 & 5 & 1 \\ 0 & -11 & -14 & -3 \\ 0 & -2 & -4 & 0 \\ 0 & -4 & -4 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 5 & 1 \\ 0 & -11 & -14 & -3 \\ 0 & 0 & -16/11 & 6/11 \\ 0 & 0 & 12/11 & 12/11 \end{vmatrix} = \\ &= \begin{vmatrix} 1 & 4 & 5 & 1 \\ 0 & -11 & -14 & -3 \\ 0 & 0 & -16/11 & 6/11 \\ 0 & 0 & 0 & 3/2 \end{vmatrix} = 1 \cdot (-11) \cdot \left(-\frac{16}{11}\right) \cdot \frac{3}{2} = 24. \end{aligned}$$

Ejercicios.

1) Demuestra que $\begin{vmatrix} 1 & 6 & 9 \\ 1 & 8 & 2 \\ 0 & 4 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 \\ 5 & 6 & 8 \\ 4 & 9 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 6 & 9 \\ 1 & 8 & 2 \\ 1 & 9 & 5 \end{vmatrix}.$

2) Calcula x en la ecuación $\begin{vmatrix} a & b & c \\ a & x & c \\ a & b & x \end{vmatrix} = 0.$

3) Calcula el valor de los siguientes determinantes:

a) $\begin{vmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{vmatrix},$ b) $\begin{vmatrix} 1+x & 1 & 1 & 1 \\ 1 & 1-x & 1 & 1 \\ 1 & 1 & 1+z & 1 \\ 1 & 1 & 1 & 1-z \end{vmatrix}.$

8.7. TIPOS DE MATRICES.

$$c) \begin{vmatrix} a & b & b & \dots & b & b \\ b & a & b & \dots & b & b \\ b & b & a & \dots & b & b \\ \vdots & \vdots & & & \vdots & \\ b & b & b & \dots & b & a \end{vmatrix}, \quad d) \begin{vmatrix} 3 & 4 & 2 & 9 & 1 \\ -1 & 5 & 4 & 6 & 2 \\ 8 & 2 & 5 & 3 & 1 \\ 2 & 2 & 0 & 6 & 4 \\ 4 & 5 & 1 & 3 & 1 \end{vmatrix}.$$

4) Calcula $A \cdot B$, $\det A$, $\det B$, $\det (AB)$ y $\det (A + B)$ para

$$a) A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 3 & 2 \\ 3 & 1 & 1 \end{pmatrix} \quad y \quad B = \begin{pmatrix} -2 & 0 & 2 \\ -1 & 1 & -1 \\ 1 & 4 & 3 \end{pmatrix};$$

$$b) A = \begin{pmatrix} 3 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} \quad y \quad B = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

- 5) ¿Es cierto que $\det (A + B) = \det A + \det B$? Dar una demostración o un contraejemplo.
- 6) ¿Es cierto que $(A + B)(A - B) = A^2 - B^2$?
- 7) Suponiendo cierta la ley $\det (A \cdot B) = (\det A) \cdot (\det B)$, prueba que $\det (A \cdot B \cdot C) = (\det A) \cdot (\det B) \cdot (\det C)$.

8.7 Tipos de matrices.

Presentamos ahora material adicional sobre matrices.

Definición. La **traspuesta** de una matriz A de orden $m \times n$, denotada por A^t , es una matriz $n \times m$ cuyos elementos son $(A^t)_{ij} = (A)_{ji}$, es decir, se obtiene cambiando filas por columnas sin alterar su orden.

La operación de hacer la traspuesta de una matriz tiene las siguientes **propiedades**:

1. La traspuesta de la traspuesta de A es la misma matriz A ; $\forall A \in M_{m \times n}$

$$(A^t)^t = A;$$

2. La traspuesta de la suma es la suma de las traspuestas; $\forall A, B \in M_{m \times n}$

$$(A + B)^t = A^t + B^t;$$

3. La traspuesta del producto es el producto intercambiado de las traspuestas;

$$\forall A \in M_{m \times n}, \forall B \in M_{n \times p}$$

$$(A B)^t = B^t A^t;$$

4. $\forall \lambda \in K, \forall A \in M_{m \times n}$

$$(\lambda A)^t = \lambda A^t;$$

5. El determinante de la traspuesta coincide con el determinante de la matriz;

$$\forall A \in M_{m \times n}$$

$$\det A^t = \det A.$$

Definición. Una matriz cuya traspuesta es ella misma se llama **simétrica**.

Para ser simétrica una matriz tiene que ser cuadrada y además los elementos simétricos respecto de la diagonal principal han de ser iguales. Por ejemplo, la siguiente es una matriz simétrica de orden 4

$$A = \begin{pmatrix} 1 & -1 & 3 & 6 \\ -1 & 2 & 5 & -2 \\ 3 & 5 & 3 & 1 \\ 6 & -2 & 1 & 4 \end{pmatrix}.$$

Definición. Una matriz A se dice que es **antisimétrica** si se verifica que $A = -A^t$.

Para ser antisimétrica una matriz tiene que ser cuadrada. los elementos simétricos respecto de la diagonal principal han de ser opuestos y los elementos de la diagonal tienen que ser nulos. Por ejemplo,

$$A = \begin{pmatrix} 0 & -1 & 3 & 6 \\ 1 & 0 & -5 & -2 \\ -3 & 5 & 0 & -1 \\ -6 & 2 & 1 & 0 \end{pmatrix}.$$

Teorema 8.4 *Toda matriz cuadrada A se puede descomponer como suma de una matriz simétrica y otra antisimétrica de forma única.*

Demostración:

Sea $A \in M_{n \times n}$. Tendrá que ser $A = S + T$ con S matriz simétrica y T matriz antisimétrica: es decir $S = S^t, T = -T^t$. Entonces

$$A^t = (S + T)^t = S^t + T^t = S - T;$$

Sumando y restando A con A^t , respectivamente

$$A + A^t = 2S \quad \implies \quad S = \frac{A + A^t}{2}, \quad A - A^t = 2T \quad \implies \quad T = \frac{A - A^t}{2}.$$

Luego,

$$A = \frac{A + A^t}{2} + \frac{A - A^t}{2} = S + T,$$

es decir, es suma de una matriz simétrica y otra antisimétrica.

c.q.d.

8.7. TIPOS DE MATRICES.

Ejemplo. Consideremos

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

entonces

$$S = \frac{\begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix}}{2} = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 0 & 3/2 \\ 3 & 3/2 & 1 \end{pmatrix}$$

y

$$T = \frac{\begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix} - \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 2 \\ 3 & 1 & 1 \end{pmatrix}}{2} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1/2 \\ 0 & -1/2 & 0 \end{pmatrix}.$$

Definición. Dada una matriz A , $m \times n$, su **matriz conjugada**, que se representa por \bar{A} , es aquella cuyos elementos son los conjugados de los elementos de A .

Por ejemplo,

$$A = \begin{pmatrix} 1 & i \\ -1 & -i \\ 1+i & -3i \end{pmatrix} \implies \bar{A} = \begin{pmatrix} 1 & -i \\ -1 & i \\ 1-i & 3i \end{pmatrix}$$

Definición. Se dice que una matriz A , $n \times n$, es **no singular** o **regular** si existe una matriz A^{-1} , $n \times n$, tal que $A \cdot A^{-1} = A^{-1} \cdot A = I$. La matriz A^{-1} se llama la **inversa** de A . Una matriz que no tiene inversa se llama **singular**.

Una manera de calcular A^{-1} es relacionarla con el determinante de la matriz y con su adjunto.

Definición. Se define el **adjunto** de una matriz A , $n \times n$, como la matriz

$$A^+ = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix},$$

donde A_{ij} son los cofactores (menores con signo) de los elementos correspondientes a_{ij} ($i, j = 1, 2, \dots, n$). [Nótese que los adjuntos de los elementos de las filas de una matriz caen en las columnas correspondientes al adjunto, es decir, se verifica la operación de trasposición].

Para encontrar la inversa de la matriz A , se dividen todos los elementos de la matriz adjunta A^+ por el valor del determinante de A :

$$A^{-1} = \frac{1}{\det A} A^+.$$

Ejemplo. Sea dada la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 3 \end{pmatrix}$, queremos encontrar su inversa.

Solución:

Empezamos hallando la matriz adjunta. Para esto necesitamos los cofactores A_{ij} :

$$\begin{aligned} A_{11} &= \begin{vmatrix} 3 & 4 \\ 4 & 3 \end{vmatrix} = -7, & A_{12} &= -\begin{vmatrix} 1 & 4 \\ 1 & 3 \end{vmatrix} = 1, & A_{13} &= \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = 1, \\ A_{21} &= -\begin{vmatrix} 2 & 3 \\ 4 & 3 \end{vmatrix} = 6, & A_{22} &= \begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} = 0, & A_{23} &= -\begin{vmatrix} 1 & 2 \\ 1 & 4 \end{vmatrix} = -2, \\ A_{31} &= \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -1, & A_{32} &= -\begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} = -1, & A_{33} &= \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = 1. \end{aligned}$$

Entonces tendremos

$$A^+ = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} -7 & 6 & -1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{pmatrix}.$$

Resulta además que, desarrollando por la primera columna

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \\ 1 & 4 & 3 \end{vmatrix} = 1 \cdot \begin{vmatrix} 3 & 4 \\ 4 & 3 \end{vmatrix} - 1 \cdot \begin{vmatrix} 2 & 3 \\ 4 & 3 \end{vmatrix} + 1 \cdot \begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix} = -7 + 6 - 1 = -2.$$

Por lo tanto A es invertible y su inversa es

$$A^{-1} = \frac{1}{\det A} A^+ = -\frac{1}{2} \begin{pmatrix} -7 & 6 & -1 \\ 1 & 0 & -1 \\ 1 & -2 & 1 \end{pmatrix} = \begin{pmatrix} -7/2 & 3 & -1/2 \\ 1/2 & 0 & -1/2 \\ 1/2 & -1 & 1/2 \end{pmatrix}.$$

Teorema 8.5 *La condición necesaria y suficiente para que una matriz tenga inversa es que su determinante sea distinto de cero.*

Demostración:

$\xrightarrow{N.}$ Si A tiene inversa, existe la matriz A^{-1} tal que $A \cdot A^{-1} = A^{-1} \cdot A = I$. Entonces $\det(A \cdot A^{-1}) = \det I$, es decir

$$\det A \cdot \det A^{-1} = \det I = 1.$$

Y esto implica que $\det A \neq 0$.

$\xleftarrow{S.}$ Si $\det A \neq 0$ podemos hallar la matriz $A^{-1} = \frac{1}{\det A} A^+$ tal que

$$A \cdot A^{-1} = A^{-1} \cdot A = I.$$

Es decir, la matriz A tiene inversa.

c.q.d.

8.7. TIPOS DE MATRICES.

A partir de la definición se siguen ciertas **propiedades** de la matriz inversa.

1. La matriz inversa de una matriz invertible es única.

Demostración:

Por ser A^{-1} la inversa de A tenemos $A \cdot A^{-1} = A^{-1} \cdot A = I$. Supongamos que exista otra matriz inversa B , luego $A \cdot B = B \cdot A = I$. Multiplicamos ahora la primera relación a la izquierda por B :

$$B \cdot (A \cdot A^{-1}) = (B \cdot A) \cdot A^{-1} = I \cdot A^{-1} = A^{-1}$$

pero el término a la izquierda coincide con la matriz B , entonces $B = A^{-1}$.

c.q.d.

2. $(A^{-1})^{-1} = A$.

Demostración:

Si A^{-1} es la inversa de A , entonces $A \cdot A^{-1} = A^{-1} \cdot A = I$. De la misma manera por ser $(A^{-1})^{-1}$ la inversa de A^{-1} tenemos $(A^{-1})^{-1} \cdot A^{-1} = A^{-1} \cdot (A^{-1})^{-1} = I$. Luego por la propiedad de la unicidad de la matriz inversa sigue que $(A^{-1})^{-1} = A$.

c.q.d.

3. $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Demostración:

Siempre por la definición de matriz inversa tenemos que

$$(A \cdot B) \cdot (A \cdot B)^{-1} = (A \cdot B)^{-1} \cdot (A \cdot B) = I.$$

Si ahora consideramos los productos $(A \cdot B) \cdot (B^{-1} \cdot A^{-1})$ y $(B^{-1} \cdot A^{-1}) \cdot (A \cdot B)$, obtenemos

$$(A \cdot B) \cdot (B^{-1} \cdot A^{-1}) = A \cdot (B \cdot B^{-1}) \cdot A^{-1} = A \cdot I \cdot A^{-1} = A \cdot A^{-1} = I,$$

$$(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = B^{-1} \cdot (A^{-1} \cdot A) \cdot B = B^{-1} \cdot I \cdot B = B^{-1} \cdot B = I.$$

Como antes, por la propiedad de la unicidad de la matriz inversa se sigue que $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

c.q.d.

4. $(A^t)^{-1} = (A^{-1})^t$.

Demostración:

Por la definición de matriz inversa, $A^t \cdot (A^t)^{-1} = (A^t)^{-1} \cdot A^t = I$. Por otro lado, usando las propiedades de la matriz traspuesta tenemos

$$A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = I^t = I \quad \text{y} \quad (A^{-1})^t \cdot A^t = (A \cdot A^{-1})^t = I^t = I.$$

Por la unicidad de la matriz inversa se sigue $(A^t)^{-1} = (A^{-1})^t$.

c.q.d.

5. El determinante de la matriz A^{-1} es el inverso del determinante de la matriz A .

Demostración:

Usando la definición de matriz inversa y las propiedades de los determinantes tenemos que

$$\det(A \cdot A^{-1}) = \det A \cdot \det A^{-1} = \det I = 1 \quad \implies \quad \det A^{-1} = \frac{1}{\det A}.$$

c.q.d.

Existen dos definiciones equivalentes de matriz **ortogonal**.

Definición 1. Se dice que una matriz cuadrada A de orden n es ortogonal si verifica que $A \cdot A^t = I$.

Definición 2. Se dice que una matriz cuadrada A de orden n es ortogonal si verifica que A es invertible y cumple que $A^{-1} = A^t$.

Veamos que las dos definiciones son equivalentes.

Def.1 \implies Def.2: Por la primera definición tenemos que $A \cdot A^t = I$, luego

$$\det(A \cdot A^t) = \det A \cdot \det A^t = \det I = 1.$$

Como $\det A = \det A^t$, sigue que $(\det A)^2 = 1$, es decir $\det A \neq 0$ que implica que la matriz A es invertible. Además tenemos que

$$A^{-1} \cdot (A \cdot A^t) = A^{-1} \cdot I = A^{-1} \quad \text{y} \quad (A^{-1} \cdot A) \cdot A^t = I \cdot A^t = A^t \quad \implies \quad A^{-1} = A^t.$$

Def.2 \implies Def.1: Inmediato, visto que

$$A \cdot A^t = A \cdot A^{-1} = I.$$

c.q.d.

Observaciones

- Si A es ortogonal entonces $(\det A)^2 = 1$ y por lo tanto $\det A = \pm 1$.
- Si una matriz es ortogonal, la suma de los productos de los elementos de la fila i de la matriz A por los correspondientes elementos de la fila j es nula si $i \neq j$ y es igual a 1, si $i = j$.

Ejemplo. La matriz $A = \begin{pmatrix} \cos x & \sen x \\ -\sen x & \cos x \end{pmatrix}$ es ortogonal, pues:

$$\begin{aligned} A \cdot A^t &= \begin{pmatrix} \cos x & \sen x \\ -\sen x & \cos x \end{pmatrix} \cdot \begin{pmatrix} \cos x & -\sen x \\ \sen x & \cos x \end{pmatrix} = \\ &= \begin{pmatrix} \cos^2 x + \sen^2 x & -\cos x \sen x + \sen x \cos x \\ -\sen x \cos x + \cos x \sen x & \sen^2 x + \cos^2 x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

8.7. TIPOS DE MATRICES.

Definición. Una matriz cuadrada U de orden n es **triangular superior** si todos los elementos situados por debajo de la diagonal principal son nulos; es decir, si $u_{ij} = 0$ para cada $i = j + 1, j + 2, \dots, n$.

Una matriz cuadrada L de orden n es **triangular inferior** si todos los elementos situados por arriba de la diagonal principal son nulos; es decir, $l_{ij} = 0$ para cada $i = 1, 2, \dots, j - 1$. Es evidente que una matriz diagonal es a la vez triangular superior e inferior.

Por ejemplo, L es una matriz triangular inferior de orden 4, mientras que U es una matriz triangular superior de orden 3

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}.$$

El cálculo del determinante de una matriz cuadrada arbitraria puede requerir un gran número de manipulaciones. Sin embargo, una matriz en forma triangular tiene un determinante fácil de calcular.

Teorema 8.6 Si $A = (a_{ij})$ es una matriz de orden n triangular superior (o triangular inferior o diagonal), entonces

$$\det A = \prod_{i=1}^n a_{ii}.$$

Definición. Una matriz $n \times n$ se llama una **matriz banda** si existen enteros p y q , tales que $1 < p, q < n$, con la propiedad de que $a_{ij} = 0$ siempre que $i + p \leq j$ ó $j + q \leq i$. El **ancho de banda** para una matriz de este tipo se define como $w = p + q - 1$.

La definición de matriz de banda fuerza a estas matrices a concentrar todos sus elementos no nulos alrededor de la diagonal. Dos casos especiales de matrices de banda que ocurren frecuentemente en la práctica son $p = q = 2$ y $p = q = 4$. Las matrices con ancho de banda 3 (que se presenta cuando $p = q = 2$) se llaman generalmente **tridiagonales** ya que tienen la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & a_{i,i-1} & a_{ii} & a_{i,i+1} & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & a_{n-2,n-1} & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_{n,n-1} & a_{nn} \end{pmatrix}.$$

Definición. Se dice que la matriz A de orden n es **estrictamente dominante diagonalmente** en el caso de que satisfaga, para cada $i = 1, 2, \dots, n$,

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|.$$

Las matrices estrictamente dominantes diagonalmente nunca son singulares.

Definición. Una matriz simétrica A de orden n se llama **positiva definida** si $\vec{x}^t A \vec{x} > 0$ para todo vector columna n -dimensional $\vec{x} \neq \vec{0}$, es decir si

$$\vec{x}^t A \vec{x} = (x_1, x_2, \dots, x_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j > 0.$$

Aquí, también, si A es una matriz positiva definida, A es no singular.

Ejercicios.

1) ¿Son ortogonales las matrices $A = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 3 & 2 \\ 1 & 5 & 6 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 3 & 5 \\ 1 & 7 & 5 \end{pmatrix}$?

2) Calcula la inversa de las siguientes matrices:

a) $A = \begin{pmatrix} 5 & 0 & 1 \\ 1 & 0 & 0 \\ 3 & 2 & -2 \end{pmatrix}$, b) $A = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 1 & 1 \\ 3 & -2 & 2 \end{pmatrix}$,

c) $A = \begin{pmatrix} 6 & 3 \\ -2 & 1 \end{pmatrix}$, d) $A = \begin{pmatrix} 5 & 1 \\ 1 & -2 \end{pmatrix}$.

3) Dadas las matrices $A = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ -2 & 1 \end{pmatrix}$ y $C = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$,

a) Efectúa, cuando sea posible, los productos AB , AC , BC y CA ;

b) Obtén las transpuestas y las inversas de A , B y C ;

c) Calcula ABC y ACB .

4) Halla la inversa de la matriz AB siendo

$$A = \begin{pmatrix} \cos x & -\operatorname{sen} x & 0 \\ \operatorname{sen} x & \cos x & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} \cos y & 0 & \operatorname{sen} y \\ 0 & 1 & 0 \\ -\operatorname{sen} y & 0 & \cos y \end{pmatrix}.$$

5) Sea A una matriz con coeficientes en un cuerpo K . Demuestra que las matrices $A \cdot A^t$ y $A^t \cdot A$ son siempre simétricas.

6) Demuestra que el producto de dos matrices ortogonales es una matriz ortogonal.

7) Verifica que la inversa de $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es $\frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

8.8 Normas de vectores y de matrices.

Sea \mathbb{R}^n el conjunto de todos los vectores columna con componentes reales, que denotaremos como $\vec{x} = (x_1, x_2, \dots, x_n)^t$ (nótese la notación de transpuesta). Para definir una distancia en \mathbb{R}^n , usaremos la idea de la **norma** de un vector, que volveremos a encontrar en la sección 9.2.

Definición. Una **norma vectorial** en \mathbb{R}^n es una función $\|\cdot\|$, de \mathbb{R}^n en \mathbb{R} con las siguientes propiedades.

- $\|\vec{x}\| \geq 0$ para todo $\vec{x} \in \mathbb{R}^n$;
- $\|\vec{x}\| = 0$ si y sólo si $\vec{x} = \vec{0}$;
- $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$ para todo $\alpha \in \mathbb{R}$ y $\vec{x} \in \mathbb{R}^n$;
- $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Definimos tres normas específicas en \mathbb{R}^n , aunque luego usaremos sólo la norma l_2 .

Definición. Las normas l_1 , l_2 y l_∞ para el vector $\vec{x} = (x_1, x_2, \dots, x_n)^t$ se definen como

$$\|\vec{x}\|_1 = \sum_{i=1}^n |x_i|, \quad \|\vec{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \quad \text{y} \quad \|\vec{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

La norma l_2 se denomina frecuentemente **norma Euclidea** del vector \vec{x} ya que representa la noción usual de distancia al origen en el caso en el que el vector \vec{x} esté en \mathbb{R} , \mathbb{R}^2 o \mathbb{R}^3 .

Ya que la norma de un vector da una medida de la distancia entre el extremo final del vector y el origen, la distancia entre dos vectores se puede definir como la norma de la diferencia de los dos vectores.

Definición. Si $\vec{x} = (x_1, x_2, \dots, x_n)^t$ e $\vec{y} = (y_1, y_2, \dots, y_n)^t$ son vectores en \mathbb{R}^n , las distancias l_1 , l_2 y l_∞ entre \vec{x} e \vec{y} se definen como

$$\|\vec{x} - \vec{y}\|_1 = \sum_{i=1}^n |x_i - y_i|,$$

$$\|\vec{x} - \vec{y}\|_2 = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad \text{y} \quad \|\vec{x} - \vec{y}\|_\infty = \max_{1 \leq i \leq n} |x_i - y_i|.$$

Teorema 8.7 Para cada $\vec{x} \in \mathbb{R}^n$,

$$\|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \sqrt{n} \|\vec{x}\|_\infty.$$

Demostración:

Sea x_j una coordenada de \vec{x} tal que $\|\vec{x}\|_\infty = \max_{1 \leq i \leq n} |x_i| = |x_j|$. Entonces

$$\|\vec{x}\|_\infty^2 = |x_j|^2 = x_j^2 \leq \sum_{i=1}^n x_i^2 \leq \sum_{i=1}^n x_j^2 = n x_j^2 = n \|\vec{x}\|_\infty^2.$$

Por lo tanto

$$\|\vec{x}\|_\infty \leq \left[\sum_{i=1}^n x_i^2 \right]^{1/2} = \|\vec{x}\|_2 \leq \sqrt{n} \|\vec{x}\|_\infty.$$

c.q.d.

A veces puede ser necesario también tener un método para medir distancias entre dos matrices $n \times n$, lo cual nuevamente requiere el uso del concepto de norma.

Definición. Una **norma matricial** en el conjunto de todas las matrices reales $n \times n$ es una función de valores reales $\|\cdot\|$, definida en este conjunto que satisface, para todas las matrices A y B , $n \times n$, y todo número real α , las siguientes propiedades:

- $\|A\| \geq 0$;
- $\|A\| = 0$ si y sólo si $A = O$;
- $\|\alpha A\| = |\alpha| \|A\|$;
- $\|A + B\| \leq \|A\| + \|B\|$;
- $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Una **distancia entre las matrices** A y B , $n \times n$, se puede definir de la manera usual como $\|A - B\|$.

Si $\|\cdot\|$ es cualquier norma vectorial en \mathbb{R}^n , entonces

$$\|A\| = \max_{\|\vec{x}\|=1} \|A\vec{x}\|$$

define una norma matricial en el conjunto de las matrices reales $n \times n$, que se llama la **norma natural**.

Consecuentemente, las normas matriciales que consideraremos tienen las formas

$$\|A\|_1 = \max_{\|\vec{x}\|_1=1} \|A\vec{x}\|_1, \quad \text{norma } l_1,$$

$$\|A\|_2 = \max_{\|\vec{x}\|_2=1} \|A\vec{x}\|_2, \quad \text{norma } l_2,$$

y

$$\|A\|_\infty = \max_{\|\vec{x}\|_\infty=1} \|A\vec{x}\|_\infty, \quad \text{norma } l_\infty.$$

8.8. NORMAS DE VECTORES Y DE MATRICES.

Teorema 8.8 Si $A = (a_{ij})$ es una matriz $n \times n$, entonces

$$a) \quad \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|;$$

$$b) \quad \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Demostración:

a) Sea \vec{x} un vector columna n -dimensional tal que su norma l_∞ sea uno; es decir, $\|\vec{x}\|_\infty = \max_{1 \leq i \leq n} |x_i| = 1$. Como $A\vec{x}$ es también un vector columna n -dimensional,

$$\begin{aligned} \|A\vec{x}\|_\infty &= \max_{1 \leq i \leq n} |(A\vec{x})_i| = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \max_{1 \leq j \leq n} |x_j| \\ &= \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \|\vec{x}\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|. \end{aligned}$$

Así que $\|A\vec{x}\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ para todo \vec{x} con $\|\vec{x}\|_\infty = 1$. Consecuentemente,

$$\|A\|_\infty = \max_{\|\vec{x}\|_\infty=1} \|A\vec{x}\|_\infty \leq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Por otro lado, si p es un entero $1 \leq p \leq n$, con

$$\sum_{j=1}^n |a_{pj}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|,$$

y \vec{x} se escoge de tal manera que

$$x_j = \begin{cases} 1, & \text{si } a_{pj} \geq 0, \\ -1, & \text{si } a_{pj} < 0, \end{cases}$$

entonces $\|\vec{x}\|_\infty = 1$ y $|a_{pj} x_j| = |a_{pj}|$ para toda $j = 1, 2, \dots, n$. Además,

$$\|A\vec{x}\|_\infty = \max_{1 \leq i \leq n} \left| \sum_{j=1}^n a_{ij} x_j \right| \geq \left| \sum_{j=1}^n a_{pj} x_j \right| = \sum_{j=1}^n |a_{pj}| = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Esto implica que

$$\|A\|_\infty = \max_{\|\vec{x}\|_\infty=1} \|A\vec{x}\|_\infty \geq \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Entonces,

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

Demostremos ahora la parte b); sea \vec{x} un vector columna n -dimensional tal que su norma l_1 sea uno; es decir, $\|\vec{x}\|_1 = \sum_{i=1}^n |x_i| = 1$. Como $A\vec{x}$ es también un vector columna n -dimensional,

$$\begin{aligned} \|A\vec{x}\|_1 &= \sum_{i=1}^n |(A\vec{x})_i| = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| = \sum_{j=1}^n \left| \left(\sum_{i=1}^n a_{ij} \right) x_j \right| \leq \\ &\leq \sum_{j=1}^n |x_j| \sum_{i=1}^n |a_{ij}| = \sum_{i=1}^n |a_{ij}| \|\vec{x}\|_1 = \sum_{i=1}^n |a_{ij}|. \end{aligned}$$

Así que $\|A\vec{x}\|_1 \leq \sum_{i=1}^n |a_{ij}|$ para toda \vec{x} con $\|\vec{x}\|_1 = 1$. Consecuentemente,

$$\|A\|_1 = \max_{\|\vec{x}\|_1=1} \|A\vec{x}\|_1 \leq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Por otro lado, si p es un entero $1 \leq p \leq n$, con

$$\sum_{i=1}^n |a_{ip}| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|,$$

y \vec{x} se escoge de tal manera que

$$x_j = \begin{cases} 1, & \text{si } j = p. \\ 0, & \text{en el resto de los casos.} \end{cases}$$

entonces $\|\vec{x}\|_1 = \sum_{j=1}^n |x_j| = 1$. Además,

$$\|A\vec{x}\|_1 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \geq \sum_{i=1}^n |a_{ip}| = \sum_{i=1}^n |a_{ip}| \sum_{j=1}^n |x_j| = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Esto implica que

$$\|A\vec{x}\|_1 = \max_{\|\vec{x}\|_1=1} \|A\vec{x}\|_1 \geq \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

Entonces,

$$\|A\vec{x}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

c.q.d.

Ejercicio.

Halla las normas l_1 , l_2 y l_{∞} para los vectores de las bases canónicas de \mathbb{R}^2 , \mathbb{R}^3 y \mathbb{R}^n .

8.9 Autovalores y autovectores.

Definición. Sea $(E, +, \cdot)$ un espacio vectorial sobre el cuerpo K . Sea $\mathcal{L}(E)$ el conjunto de endomorfismos de E . Dado $f \in \mathcal{L}(E)$, se dice que un vector $\vec{x} \neq \vec{0}$ y $\vec{x} \in E$ es un **vector propio** o **autovector**, si existe un escalar $\lambda \in K$ tal que $f(\vec{x}) = \lambda \vec{x}$. Al valor λ se le llama **valor propio** o **autovalor** asociado al autovector \vec{x} .

Definición. Se llama valor propio del endomorfismo $f: E \rightarrow E$ a todo escalar $\lambda \in K$ tal que existe al menos un vector $\vec{x} \neq \vec{0}$ que cumple que $f(\vec{x}) = \lambda \vec{x}$.

Si por ejemplo consideramos

$$f(\vec{x}_1) = -3\vec{x}_1, \quad f(\vec{x}_2) = -3\vec{x}_2, \quad f(\vec{x}_3) = -3\vec{x}_3,$$

al autovalor $\lambda = -3$, están asociados tres autovectores \vec{x}_1 , \vec{x}_2 y \vec{x}_3 .

Ejemplo. Sea f un endomorfismo de \mathbb{R}^2 cuya matriz respecto a la base canónica es

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

¿tiene dicho endomorfismo autovectores?

Solución:

Si $\vec{x} = (x_1, x_2) \neq (0, 0)$ fuese un autovector, existiría un escalar $\lambda \in \mathbb{R}$ tal que $f(\vec{x}) = \lambda \vec{x}$; es decir,

$$f(\vec{x}) = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \vec{x}.$$

Este sistema matricial equivale al sistema de ecuaciones

$$\begin{cases} x_1 - 2x_2 = \lambda x_1 \\ 2x_1 + x_2 = \lambda x_2 \end{cases} \Rightarrow \begin{cases} (1 - \lambda)x_1 - 2x_2 = 0 \\ 2x_1 + (1 - \lambda)x_2 = 0 \end{cases}$$

que es un sistema homogéneo cuya única solución es $x_1 = x_2 = 0$, dado que el determinante de su matriz de coeficientes vale

$$\begin{vmatrix} (1 - \lambda) & -2 \\ 2 & (1 - \lambda) \end{vmatrix} = (1 - \lambda)^2 + 4 \neq 0 \quad \forall \lambda \in \mathbb{R}.$$

Luego el endomorfismo no tiene autovectores no nulos.

Ejemplo. Sea E el espacio vectorial real formado por las funciones reales $\phi(x)$ que admiten derivadas de cualquier orden. La aplicación $f: E \rightarrow E$ definida por $f(\phi(x)) = \phi'(x)$ es una aplicación lineal y por lo tanto es un endomorfismo.

- la función e^x es un autovector, siendo su autovalor $\lambda = 1$, pues $f(e^x) = [e^x]' = e^x$;
- la función e^{2x} es un autovector, siendo su autovalor $\lambda = 2$, pues $f(e^{2x}) = [e^{2x}]' = 2e^{2x}$;

- la función $\operatorname{sen} x$ no es un autovector puesto que $f(\operatorname{sen} x) = [\operatorname{sen} x]' = \cos x$ y $\cos x \neq \lambda \operatorname{sen} x$.

Pasamos a estudiar algunas **propiedades** de los autovectores.

1. Un autovector \vec{x} no puede corresponder a dos autovalores distintos λ y μ .

Demostración:

Si así fuese

$$\begin{cases} f(\vec{x}) = \lambda \vec{x} \\ f(\vec{x}) = \mu \vec{x} \end{cases} \Rightarrow \lambda \vec{x} = \mu \vec{x}$$

y como $\vec{x} \neq \vec{0}$, por las propiedades de los espacios vectoriales tiene que ser $\lambda = \mu$.

c.q.d.

2. Si \vec{x} es un autovector del endomorfismo f correspondiente al autovalor λ , entonces cualquier vector $\mu \vec{x}$, donde $\mu \in K$, es también autovector asociado al autovalor λ .

Demostración:

Como \vec{x} es autovector asociado al autovalor λ tenemos que $f(\vec{x}) = \lambda \vec{x}$; además, siendo f un endomorfismo y por las propiedades de los espacios vectoriales, tenemos

$$f(\mu \vec{x}) = \mu f(\vec{x}) = \mu (\lambda \vec{x}) = \lambda (\mu \vec{x}).$$

c.q.d.

Definición. Un subespacio vectorial F de un espacio vectorial E , se dice **invariante** por el endomorfismo $f: E \rightarrow E$, si $f(F) \subset F$, es decir, si la imagen del subespacio F está incluida en F .

3. Si F es un subespacio vectorial de E de dimensión uno e invariante por el endomorfismo f , entonces cualquier vector no nulo de F es autovector.

Demostración:

Si $\vec{x} \in F$ y $\vec{x} \neq \vec{0}$, siendo F de dimensión uno, se sigue que $\{\vec{x}\}$ es una base de F . Como F es invariante, por la definición, $f(\vec{x}) \in F$, luego $f(\vec{x})$ se podrá expresar en función de la base $\{\vec{x}\}$ de F , es decir $f(\vec{x}) = \alpha \vec{x}$, por lo tanto \vec{x} es autovector correspondiente al autovalor α .

c.q.d.

4. Cuando \vec{x} es un autovector, el espacio vectorial engendrado por \vec{x} es invariante por el endomorfismo f .

8.9. AUTOVALORES Y AUTOVECTORES.

Demostración:

Esto sigue de la propiedad 2. Si \vec{x} es un autovector, sabemos que $f(\vec{x}) = \lambda \vec{x}$; además el subespacio engendrado por \vec{x} será $F = \{\mu \vec{x} / \mu \in \mathbb{R}\}$. Luego un vector cualquiera del subespacio F será del tipo $\mu \vec{x}$, y

$$f(\mu \vec{x}) = \lambda (\mu \vec{x}) = (\lambda \mu) \vec{x} = \mu_1 \vec{x} \in F,$$

como la imagen de cualquier vector de F pertenece a F , tenemos que $f(F) \subset F$.

c.q.d.

Podemos resumir las propiedades 2, 3 y 4 diciendo que los autovectores del endomorfismo f son los vectores no nulos de los subespacios invariantes de dimensión uno.

Teorema 8.9 *La condición necesaria y suficiente para que λ sea autovalor de f es que el endomorfismo $(f - \lambda i)$ no sea inyectivo.*

(Nótese que $i: E \rightarrow E$ es el endomorfismo identidad).

Demostración:

\xrightarrow{N} Si λ es autovalor de f , existe $\vec{x} \in E$ tal que $f(\vec{x}) = \lambda \vec{x}$, pero como $i(\vec{x}) = \vec{x}$, podemos escribir

$$f(\vec{x}) = \lambda i(\vec{x}) \Leftrightarrow f(\vec{x}) - \lambda i(\vec{x}) = \vec{0} \Leftrightarrow (f - \lambda i)(\vec{x}) = \vec{0}.$$

Esto quiere decir que $\vec{x} \in \text{Ker}(f - \lambda i)$. Siendo $\vec{x} \neq \vec{0}$, se sigue que $\text{Ker}(f - \lambda i) \neq \{\vec{0}\}$ y entonces por el teorema 8.2 el endomorfismo $(f - \lambda i)$ no es inyectivo.

\xleftarrow{S} Si $(f - \lambda i)$ no es inyectivo, por el teorema 8.2 se sigue: $\text{Ker}(f - \lambda i) \neq \{\vec{0}\}$, es decir, existe un vector $\vec{x} \neq \vec{0}$ tal que

$$(f - \lambda i)(\vec{x}) = \vec{0} \Leftrightarrow f(\vec{x}) - \lambda i(\vec{x}) = \vec{0} \Leftrightarrow f(\vec{x}) = \lambda i(\vec{x}) = \lambda \vec{x},$$

es decir, λ es autovalor de f .

c.q.d.

Pasamos ahora a definir el polinomio característico.

Sea E un espacio vectorial de dimensión n , sobre el cuerpo K . Sea $f: E \rightarrow E$ un endomorfismo de E y $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ una base de E . Como vimos en la sección 8.3, hay una matriz A asociada al endomorfismo f en la base B ,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Además, en la base B tenemos

$$\begin{aligned}\lambda i(\vec{u}_1) &= \lambda \vec{u}_1 + 0 \vec{u}_2 + \dots + 0 \vec{u}_n = \lambda \vec{u}_1 \\ \lambda i(\vec{u}_2) &= 0 \vec{u}_1 + \lambda \vec{u}_2 + \dots + 0 \vec{u}_n = \lambda \vec{u}_2 \\ &\vdots \\ \lambda i(\vec{u}_n) &= 0 \vec{u}_1 + 0 \vec{u}_2 + \dots + \lambda \vec{u}_n = \lambda \vec{u}_n\end{aligned}$$

es decir, la matriz asociada al endomorfismo λi es

$$\lambda I_n = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Entonces, la matriz asociada al endomorfismo $(f - \lambda i)$ es

$$A - \lambda I_n = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Definición. Cuando λ recorre el cuerpo K , el determinante $|A - \lambda I_n|$ es un polinomio en λ de grado n llamado **polinomio característico** de A , y se denota con

$$p(\lambda) = \det(A - \lambda I_n) = |A - \lambda I_n|.$$

Teorema 8.10 *El polinomio característico es independiente de las bases elegidas en el espacio vectorial E .*

Demostración:

Sea $B_1 = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ una base de E . Sea A la matriz asociada al endomorfismo $f: E \rightarrow E$ en la base B_1 de E . El polinomio característico es $p_1(\lambda) = \det(A - \lambda I_n)$.

Sea $B_2 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ otra base de E . Sea B la matriz asociada al endomorfismo $f: E \rightarrow E$ en la nueva base B_2 de E . El polinomio característico es $p_2(\lambda) = \det(B - \lambda I_n)$.

Por lo que vimos en la sección 8.3, sabemos que existe una matriz P , llamada matriz de cambio, tal que $B = P^{-1} A P$, luego

$$B - \lambda I_n = P^{-1} A P - \lambda I_n = P^{-1} A P - \lambda P^{-1} I_n P = P^{-1} (A - \lambda I_n) P$$

y entonces

$$\begin{aligned}|B - \lambda I_n| &= |P^{-1} (A - \lambda I_n) P| = |P^{-1}| \cdot |A - \lambda I_n| \cdot |P| = \\ &= |A - \lambda I_n| \cdot |P^{-1}| \cdot |P| = |A - \lambda I_n| \cdot |P^{-1} P| = \\ &= |A - \lambda I_n| \cdot |I_n| = |A - \lambda I_n|.\end{aligned}$$

c.q.d.

8.9. AUTOVALORES Y AUTOVECTORES.

Teorema 8.11 *El valor λ es autovalor del endomorfismo f si y sólo si $|A - \lambda I| = 0$.*

Demostración:

\xrightarrow{N} En el teorema 8.9 hemos visto que la condición necesaria y suficiente para que λ sea autovalor de f es que el endomorfismo $(f - \lambda i)$ no sea inyectivo, es decir que $\text{Ker}(f - \lambda i) \neq \{\vec{0}\}$, o sea existe un vector $\vec{x} \neq \vec{0}$ tal que $(f - \lambda i)(\vec{x}) = \vec{0}$. Pasando a la matriz que representa el endomorfismo, tenemos $(A - \lambda I)\vec{x} = \vec{0}$, o que es lo mismo

$$(A - \lambda I)\vec{x} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \vec{0}.$$

Esta última ecuación matricial representa un sistema homogéneo de ecuaciones lineales, que, como $\vec{x} = (x_1, x_2, \dots, x_n) \neq \vec{0}$, admite soluciones distintas de la trivial, y por lo tanto el determinante de la matriz de coeficientes del sistema tiene que ser nulo, es decir

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Pero este determinante coincide con el polinomio característico.

\xleftarrow{S} Por otro lado, si $|A - \lambda I| = 0$, entonces el sistema homogéneo de ecuaciones lineales que tiene por coeficientes los elementos de la matriz $A - \lambda I$ admite soluciones distintas de la trivial, es decir, existe $\vec{x} \neq \vec{0}$ tal que $(f - \lambda i)(\vec{x}) = \vec{0}$, que es lo mismo que $\text{Ker}(f - \lambda i) \neq \{\vec{0}\}$; entonces por el 8.9 el endomorfismo $(f - \lambda i)$ no es inyectivo, y λ es un autovalor.

c.q.d.

Podemos resumir diciendo que los autovalores son las raíces de la ecuación característica, es decir la ecuación obtenida igualando a cero el polinomio característico. En total habrá n autovalores. Si las raíces son distintas se llaman autovalores simples; si algunas raíces son iguales, se llaman autovalores múltiples, y tendrán un orden de multiplicidad determinado.

Ejemplo 1. Suponed que al resolver un polinomio característico obtenemos

$$\lambda_1 = 1, \quad \lambda_2 = 3, \quad \lambda_3 = 1, \quad \lambda_4 = 1, \quad \lambda_5 = 0, \quad \lambda_6 = 1.$$

El cero y el tres son autovalores simples (se repiten una sola vez), mientras que el autovalor uno es múltiple, con multiplicidad cuatro.

Ejemplo 2. Supongamos que

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

y que un autovalor sea $\lambda = 5$. Calcular el autovector asociado a dicho autovalor.

Solución:

Para calcular un autovector asociado a dicho autovalor, construimos, para el valor $\lambda = 5$, el sistema $(A - \lambda I) \vec{x} = \vec{0}$, es decir

$$\begin{pmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2-5 & 1 & 1 \\ 2 & 3-5 & 1 \\ 1 & 1 & 2-5 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -3 & 1 & 1 \\ 2 & -2 & 1 \\ 1 & 1 & -3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} -3x_1 + x_2 + x_3 = 0 \\ 2x_1 - 2x_2 + x_3 = 0 \\ x_1 + x_2 - 3x_3 = 0 \end{cases}$$

Resolviendo este sistema obtenemos $x_1 = x_3$ y $x_2 = 2x_3$, es decir el conjunto de autovectores para $\lambda = 5$ es $\{(x_3, 2x_3, x_3)\}$, o también podemos decir que $\text{Ker}(f - 5i) = \{(x_3, 2x_3, x_3)\}$. Uno en particular es el autovector $(1, 2, 1)$.

Ejemplo 3. Sea

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

Determina los autovalores y autovectores.

Solución:

Para calcular los autovalores de A consideremos

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 0 & 1 \\ 2 & 2-\lambda & 1 \\ -1 & 0 & -\lambda \end{pmatrix} = \\ &= (1-\lambda) \cdot \det \begin{pmatrix} 2-\lambda & 1 \\ 0 & -\lambda \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 2 & 2-\lambda \\ -1 & 0 \end{pmatrix} = \\ &= (1-\lambda)(2-\lambda)(-\lambda) + (2-\lambda) = (2-\lambda)(\lambda^2 - \lambda + 1). \end{aligned}$$

Los autovalores de A son las soluciones de $p(\lambda) = 0$, es decir,

$$\lambda_1 = 2, \quad \lambda_2 = \frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad \lambda_3 = \frac{1}{2} - \frac{\sqrt{3}}{2}i.$$

Un autovector \vec{x} de A asociado con λ_1 es una solución del sistema $(A - \lambda I) \vec{x} = \vec{0}$:

$$\begin{pmatrix} -1 & 0 & 1 \\ 2 & 0 & 1 \\ -1 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Por lo tanto,

$$\begin{aligned} -x_1 + x_3 &= 0, \\ 2x_1 + x_3 &= 0, \\ -x_1 - 2x_3 &= 0, \end{aligned}$$

8.9. AUTOVALORES Y AUTOVECTORES.

sistema que tiene por solución a $x_1 = x_3 = 0$ y x_2 arbitrario. En particular, $\vec{x} = (0, 1, 0)^t$ es un autovector de A correspondiente al autovalor $\lambda_1 = 2$.

Definición. Sea λ un autovalor de un endomorfismo f de un espacio vectorial E definido sobre el cuerpo K . Definimos E_λ como el conjunto de los autovectores asociados a λ , junto con el vector nulo, $\vec{0}$. A E_λ se le llama **espacio propio de λ** .

Teorema 8.12 *El espacio propio de λ , E_λ , es un subespacio vectorial de E .*

Demostración:

Por la definición tenemos

$$E_\lambda = \{\vec{x} \in E / f(\vec{x}) = \lambda \vec{x}\} \cup \{\vec{0}\} = \text{Ker}(f - \lambda i).$$

Para todo $\alpha, \beta \in K$ y para todo $\vec{x}, \vec{y} \in E_\lambda$ se tiene que cumplir $\alpha \vec{x} + \beta \vec{y} \in E_\lambda$. Esto es cierto por ser f un endomorfismo:

$$f(\alpha \vec{x} + \beta \vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}) = \alpha (\lambda \vec{x}) + \beta (\lambda \vec{y}) = \lambda (\alpha \vec{x} + \beta \vec{y})$$

es decir, $\alpha \vec{x} + \beta \vec{y}$ es un autovector con autovalor λ , y por lo tanto pertenece a E_λ .

c.q.d.

Teorema 8.13 Propiedad fundamental de los autovalores.

Si $\lambda_1, \lambda_2, \dots, \lambda_p$ son autovalores simples (es decir distintos dos a dos) de un endomorfismo f de un espacio vectorial E , entonces los autovectores correspondientes $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ son linealmente independientes.

Demostración:

Hagamos la demostración por reducción al absurdo. Supongamos que $\lambda_i \neq \lambda_j$ siendo $i \neq j$ y que los vectores $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k, \vec{x}_{k+1}, \dots, \vec{x}_p\}$ son linealmente dependientes. Supongamos entonces que los primeros k vectores $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k\}$ sean linealmente independientes. Entonces, los vectores $\vec{x}_{k+1}, \dots, \vec{x}_p$ son combinación lineal de $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k$, es decir

$$\vec{x}_i = \alpha_{i1} \vec{x}_1 + \alpha_{i2} \vec{x}_2 + \dots + \alpha_{ik} \vec{x}_k = \sum_{j=1}^k \alpha_{ij} \vec{x}_j, \quad i = k+1, \dots, p.$$

Sean ahora λ_i los autovalores correspondientes a los autovectores \vec{x}_i , con $i = k+1, \dots, p$; y λ_j los autovalores correspondientes a los autovectores \vec{x}_j , con $j = 1, \dots, k$. Entonces,

$$f(\vec{x}_i) = \lambda_i \vec{x}_i = \lambda_i \left(\sum_{j=1}^k \alpha_{ij} \vec{x}_j \right) = \sum_{j=1}^k \lambda_i \alpha_{ij} \vec{x}_j,$$

$$f(\vec{x}_i) = f \left(\sum_{j=1}^k \alpha_{ij} \vec{x}_j \right) = \sum_{j=1}^k \alpha_{ij} f(\vec{x}_j) = \sum_{j=1}^k \alpha_{ij} \lambda_j \vec{x}_j.$$

Los dos sumatorios de la derecha tienen que coincidir, luego

$$\sum_{j=1}^k \lambda_i \alpha_{ij} \vec{x}_j = \sum_{j=1}^k \alpha_{ij} \lambda_j \vec{x}_j \Rightarrow \sum_{j=1}^k \lambda_i \alpha_{ij} \vec{x}_j - \alpha_{ij} \lambda_j \vec{x}_j = \sum_{j=1}^k \alpha_{ij} (\lambda_i - \lambda_j) \vec{x}_j = \vec{0}.$$

Hemos obtenido una combinación lineal de los vectores linealmente independientes $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k$ que nos da el vector nulo. Entonces, sus coeficientes han de ser nulos, es decir

$$\alpha_{ij} (\lambda_i - \lambda_j) = 0, \quad \forall i, j.$$

Si aquí fuese $\alpha_{ij} = 0$, esto implicaría que $\vec{x}_i = \vec{0}$. Pero esto no puede suceder porque \vec{x}_i es un autovector y por tanto no puede ser nulo. Luego $\alpha_{ij} \neq 0$. Entonces tiene que ocurrir que $\lambda_i - \lambda_j = 0$, esto es

$$\lambda_i = \lambda_j, \quad i = k+1, \dots, p, \quad j = 1, \dots, k.$$

Y esta es una contradicción visto que partiamos de la hipótesis de que $\lambda_i \neq \lambda_j$, para $i \neq j$. Entonces los vectores $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k, \vec{x}_{k+1}, \dots, \vec{x}_p\}$ son linealmente independientes.

c.q.d.

Nota. Todo el estudio hecho sobre los autovalores y autovectores de un endomorfismo se traslada de forma instantánea a los autovalores y autovectores de matrices cuadradas.

Definición. Dos matrices A y B se dice que son **semejantes**, si existe una matriz P , no singular (esto es con determinante no nulo), tal que $B = P^{-1}AP$.

Claramente dos matrices semejantes tienen el mismo polinomio característico (la demostración se hace de la misma forma que demostramos el teorema 8.10). Luego, dos matrices semejantes tienen los mismos autovalores y con igual orden.

El recíproco de estas propiedades no es, en general, cierto; es decir, dos matrices pueden tener los mismos autovalores, con igual orden, y sin embargo no ser semejantes. Veamos un ejemplo:

Sean las matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Las dos tienen por ecuación característica $(\lambda - 1)^2 = 0$ y por lo tanto las dos matrices tiene el autovalor $\lambda = 1$, de orden 2. Sin embargo, no son semejantes, pues, de serlo, existiría una matriz regular de orden 2

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{tal que} \quad B = P^{-1}AP \quad \Leftrightarrow \quad PB = AP$$

esto es

$$PB = \begin{pmatrix} a & a+b \\ c & c+d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = AP \quad \Leftrightarrow \quad \begin{cases} a = 0, \\ c = 0. \end{cases}$$

con lo cual la matriz P sería singular.

8.10. DIAGONALIZACIÓN DE MATRICES.

Ejercicios.

- 1) Dado el endomorfismo $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ cuya matriz asociada es

$$\begin{pmatrix} 2 & -2 & 1 \\ 2 & -8 & -2 \\ 1 & 2 & 2 \end{pmatrix}$$

calcula los autovalores y autovectores.

- 2) Dado el endomorfismo $f: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ cuya matriz asociada es

$$\begin{pmatrix} 0 & 1 & 5 & 9 \\ 2 & 1 & 6 & 8 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

calcula los autovalores y autovectores.

8.10 Diagonalización de matrices.

Definición. Un endomorfismo $f: E \rightarrow E$, donde E es un espacio vectorial de dimensión finita n sobre un cuerpo K , se dice que es **diagonalizable**, si existe una base de E en la que la matriz asociada a f es una matriz diagonal. Diagonalizar f es determinar una base, en caso de que la haya, en la que su matriz asociada sea diagonal, y calcular dicha matriz.

Definición. Dada una matriz A de orden n , diremos que es **diagonalizable**, si existe una matriz P no singular tal que $D = P^{-1}AP$, siendo D una matriz diagonal. A la matriz D se le llama **forma diagonal** de la matriz A .

Como una matriz A puede considerarse como la matriz asociada a un endomorfismo f , vamos a estudiar la diagonalización de endomorfismos y matrices de forma paralela.

Consideremos dos casos:

- Si los autovalores son simples.

Supongamos que $\lambda_1, \lambda_2, \dots, \lambda_n$ sean autovalores de f distintos dos a dos. Esto implica que los autovectores correspondientes $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ son linealmente independientes, por el teorema 8.13. Como $\dim E = n$, los autovectores $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\}$ forman una base. Además

$$\begin{aligned} f(\vec{x}_1) &= \lambda \vec{x}_1 + 0 \vec{x}_2 + \dots + 0 \vec{x}_n = \lambda \vec{x}_1 \\ f(\vec{x}_2) &= 0 \vec{x}_1 + \lambda \vec{x}_2 + \dots + 0 \vec{x}_n = \lambda \vec{x}_2 \\ &\vdots \\ f(\vec{x}_n) &= 0 \vec{x}_1 + 0 \vec{x}_2 + \dots + \lambda \vec{x}_n = \lambda \vec{x}_n \end{aligned}$$

luego la matriz asociada a la base $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ es la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Resumiendo, la matriz A era la matriz asociada a f cuando en E nos daban una base determinada. Al trabajar con la base formada por los autovectores, la matriz asociada es D , luego por lo visto en la sección 8.3 existe una matriz P , matriz cambio de base, regular, tal que $D = P^{-1} A P$; luego la matriz A es diagonalizable. Como veremos en los ejemplos, la matriz de cambio P tiene por columnas las coordenadas de los autovectores.

- Si los autovalores son múltiples.

Dado el endomorfismo $f: E \rightarrow E$, supongamos que los autovalores sean $\lambda_1, \lambda_2, \dots, \lambda_p$ cada uno con multiplicidad k_1, k_2, \dots, k_p , con $k_1 + k_2 + \dots + k_p = n$. Sea A la matriz asociada a f en una determinada base. Se puede demostrar el siguiente resultado

Teorema 8.14

$$A \text{ es diagonalizable} \iff \dim \text{Ker}(f - \lambda_j i) = k_j.$$

Resulta entonces que no todas las matrices cuadradas (de elementos complejos en general) son diagonalizables.

Sin embargo, las matrices reales simétricas (que aparecen con bastante frecuencia en problemas de Física), tienen algunas propiedades importantes por sus aplicaciones.

1. Los autovalores de una matriz real y simétrica A son todos reales.

Demostración:

Por ser la matriz A de elementos reales, la ecuación característica $|A - \lambda I| = 0$ es una ecuación de coeficientes reales, y por tanto, si tiene una raíz imaginaria $\lambda_1 = \alpha + \beta i$, tendrá también la raíz imaginaria conjugada $\lambda_2 = \alpha - \beta i$. Sean \bar{x}_1 y \bar{x}_2 los autovectores asociados a λ_1 y λ_2 , respectivamente. Con los mismos símbolos \bar{x}_1 y \bar{x}_2 denotemos las matrices columnas formadas por las coordenadas del los autovectores. Además, por ser simétrica, $A = A^t$. Luego, tendremos que

$$\begin{aligned} \begin{cases} A \bar{x}_1 = \lambda_1 \bar{x}_1 \\ A \bar{x}_2 = \lambda_2 \bar{x}_2 \end{cases} &\Rightarrow \begin{cases} \bar{x}_2^t A \bar{x}_1 = \lambda_1 \bar{x}_2^t \bar{x}_1 \\ \bar{x}_1^t A \bar{x}_2 = \lambda_2 \bar{x}_1^t \bar{x}_2 \end{cases} \Rightarrow \begin{cases} (\bar{x}_2^t A \bar{x}_1)^t = (\lambda_1 \bar{x}_2^t \bar{x}_1)^t \\ \bar{x}_1^t A \bar{x}_2 = \lambda_2 \bar{x}_1^t \bar{x}_2 \end{cases} \\ &\Rightarrow \begin{cases} (A \bar{x}_1)^t \bar{x}_2 = \lambda_1 \bar{x}_1^t \bar{x}_2 \\ \bar{x}_1^t A \bar{x}_2 = \lambda_2 \bar{x}_1^t \bar{x}_2 \end{cases} \Rightarrow \begin{cases} \bar{x}_1^t A^t \bar{x}_2 = \lambda_1 \bar{x}_1^t \bar{x}_2 \\ \bar{x}_1^t A \bar{x}_2 = \lambda_2 \bar{x}_1^t \bar{x}_2 \end{cases} \Rightarrow \end{aligned}$$

8.10. DIAGONALIZACIÓN DE MATRICES.

$$\begin{cases} \vec{x}_1^t A \vec{x}_2 = \lambda_1 \vec{x}_1^t \vec{x}_2 \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \Rightarrow \lambda_1 = \lambda_2 \Rightarrow \alpha + \beta i = \alpha - \beta i \Rightarrow \\ \beta = 0 \Rightarrow \lambda_1 \text{ es real.}$$

c.q.d.

2. Si la matriz A es real y simétrica y los autovalores son todos distintos, entonces los autovectores son ortogonales dos a dos.

Demostración:

Supongamos $\lambda_1 \neq \lambda_2$, entonces como en la propiedad anterior tenemos

$$\begin{cases} A \vec{x}_1 = \lambda_1 \vec{x}_1 \\ A \vec{x}_2 = \lambda_2 \vec{x}_2 \end{cases} \Rightarrow \begin{cases} \vec{x}_2^t A \vec{x}_1 = \lambda_1 \vec{x}_2^t \vec{x}_1 \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \Rightarrow \begin{cases} (\vec{x}_2^t A \vec{x}_1)^t = (\lambda_1 \vec{x}_2^t \vec{x}_1)^t \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \\ \Rightarrow \begin{cases} (A \vec{x}_1)^t \vec{x}_2 = \lambda_1 \vec{x}_1^t \vec{x}_2 \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \Rightarrow \begin{cases} \vec{x}_1^t A^t \vec{x}_2 = \lambda_1 \vec{x}_1^t \vec{x}_2 \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \Rightarrow \\ \begin{cases} \vec{x}_1^t A \vec{x}_2 = \lambda_1 \vec{x}_1^t \vec{x}_2 \\ \vec{x}_1^t A \vec{x}_2 = \lambda_2 \vec{x}_1^t \vec{x}_2 \end{cases} \text{ restando } \Rightarrow 0 = (\lambda_1 - \lambda_2) \vec{x}_1^t \vec{x}_2$$

siendo $\lambda_1 \neq \lambda_2$ tendrá que ser

$$\Rightarrow \vec{x}_1^t \vec{x}_2 = 0 \Rightarrow (x_{11}, x_{21}, \dots, x_{n1}) \cdot \begin{pmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{n2} \end{pmatrix} = 0,$$

como el producto escalar de los dos vectores es nulo los vectores son ortogonales.

c.q.d.

Como consecuencia de las dos propiedades tenemos que

Teorema 8.15 Si la matriz es real y simétrica y los autovalores son todos distintos entonces los autovectores forman una base ortonormal. Además, la matriz de paso P , que tiene por columnas las componentes de los autovectores, es una matriz ortogonal, es decir $P^{-1} = P^t$.

(La segunda parte se puede demostrar fácilmente, haciendo las cuentas).

Estas últimas dos propiedades se cumplen para toda matriz real y simétrica; es decir,

Teorema 8.16 Toda matriz real y simétrica es diagonalizable cualquiera que sea el orden de multiplicidad de sus autovalores. La base formada por los autovectores divididos cada uno por su norma es una base ortonormal. La matriz de paso P que permite la diagonalización $D = P^{-1} A P$, es una matriz ortogonal.

Ejemplo. Diagonaliza la matriz cuadrada

$$A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 7 & 3 \\ 0 & 3 & 7 \end{pmatrix},$$

y calcula la base ortonormal en la cual la matriz asociada es diagonal.

Solución:

Empecemos calculando los autovalores.

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 4 - \lambda & 0 & 0 \\ 0 & 7 - \lambda & 3 \\ 0 & 3 & 7 - \lambda \end{vmatrix} = (4 - \lambda)(7 - \lambda)^2 - 9(4 - \lambda) = \\ &= (4 - \lambda)(49 - 14\lambda + \lambda^2 - 9) = \\ &= (4 - \lambda)(\lambda^2 - 14\lambda + 40) = \\ &= (4 - \lambda)(\lambda - 10)(\lambda - 4) \end{aligned}$$

es decir, los autovalores son

$$\lambda_1 = 10 \text{ simple}, \quad \lambda_2 = 4 \text{ doble.}$$

Pasamos ahora a calcular los autovectores.

Para $\lambda_1 = 10$ tenemos

$$\begin{aligned} (A - \lambda I) \vec{x} = \vec{0} &\Leftrightarrow \begin{pmatrix} 4 - \lambda & 0 & 0 \\ 0 & 7 - \lambda & 3 \\ 0 & 3 & 7 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \\ \begin{pmatrix} -6 & 0 & 0 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} &\Leftrightarrow \begin{cases} -6x_1 = 0 \\ -3x_2 + 3x_3 = 0 \\ 3x_2 - 3x_3 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 0, \\ x_2 = x_3. \end{cases} \end{aligned}$$

Entonces, para $\lambda_1 = 10$ podemos escoger $\vec{x} = (0, x_2, x_2)$.

Para $\lambda_2 = 4$ tenemos

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 3x_2 + 3x_3 = 0 \\ 3x_2 + 3x_3 = 0 \end{cases} \Leftrightarrow x_3 = -x_2.$$

Entonces, para $\lambda_2 = 4$ podemos escoger $\vec{x} = (x_1, x_2, -x_2)$.

Queremos ahora formar una base ortogonal: para el primer autovector consideremos $(0, 1, 1)$, y para el segundo tenemos que tomar dos vectores linealmente independientes y ortogonales entre sí. Sea el primero de ellos el vector $(1, 1, -1)$, busquemos el valor de x_2 tal que el segundo escogido como $(-1, x_2, -x_2)$ sea ortogonal al primero:

$$(1, 1, -1) \cdot (-1, x_2, -x_2) = 0 \Leftrightarrow -1 + x_2 + x_2 = 0 \Leftrightarrow x_2 = \frac{1}{2}.$$

8.10. DIAGONALIZACIÓN DE MATRICES.

Ahora normalizamos los vectores

$$\|(0, 1, 1)\| = \sqrt{1+1} = \sqrt{2}, \quad \|(1, 1, -1)\| = \sqrt{1+1+1} = \sqrt{3},$$

$$\left\| \left(-1, \frac{1}{2}, \frac{1}{2}\right) \right\| = \sqrt{1 + \frac{1}{4} + \frac{1}{4}} = \frac{\sqrt{6}}{2}.$$

Por lo tanto, una base ortonormal que diagonaliza la matriz A es

$$\left\{ \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right), \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right) \right\},$$

y la matriz de paso es

$$P = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{pmatrix}.$$

Se puede demostrar que $P^t A P = D$, es decir que

$$\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 0 \\ 0 & 7 & 3 \\ 0 & 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Ejercicios.

- 1) Encuentra la condición necesaria y suficiente para que la matriz de coeficientes reales

$$\begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$$

sea diagonalizable.

- 2) Dada la matriz

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & b \\ 3 & 0 & a \end{pmatrix}$$

estudia para qué valores de a y b es diagonalizable y calcula los autovectores correspondientes en el caso de que lo sea.

- 3) Sea A una matriz de orden 2 con coeficiente reales. Halla la condición necesaria y suficiente para que los autovectores sean iguales.

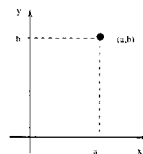
Capítulo 9

La geometría del espacio euclídeo.

9.1 Vectores en el espacio tridimensional.

Definición. Los puntos P en el plano se representan con pares ordenados (a, b) ; los números reales a y b se llaman **coordenadas cartesianas** del punto P .

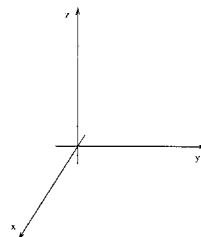
Si trazamos dos rectas perpendiculares que se intersectan en un punto O , que llamamos **origen**, las dos rectas son los **ejes x, y** . A partir de ese origen O se miden distancias sobre los ejes, y después de escoger unidades en esos ejes, podemos determinar las dos distancias a y b (trazando perpendiculares de P a los ejes). En particular, a se llama la componente x de P , o **abscisa**, y b es la componente y , o **ordenada**.



Esta es la representación de los puntos en el plano o \mathbb{R}^2 ; en el espacio o \mathbb{R}^3 la representación de los puntos es análoga, mediante ternas ordenadas de números reales.

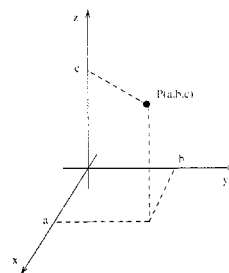
Para establecer un sistema de coordenadas cartesianas en el espacio tridimensional partimos de un sistema de coordenadas cartesianas en el plano, o sea Oxy . Por el punto O , que seguimos llamando origen, trazamos una recta que sea perpendicular a las otras dos. A esta tercera recta la llamamos eje z , usando la misma escala que en las otras dos y asignando la coordenada $z = 0$ al origen.

Regla de la mano derecha. Por conveniencia, orientamos el eje z de tal manera que $Oxyz$ forme un sistema **orientado positivamente**; es decir, que si el índice de la mano derecha apunta en la dirección de las x positivas y el dedo medio de la misma mano apunta en la dirección de las y positivas, el pulgar señalará en la dirección de las z positivas.



Definición. Los puntos P en el espacio se representan con ternas ordenadas (a, b, c) ; los números reales a, b y c se llaman **coordenadas cartesianas** del punto P . En particular, a se llama la componente x (o primera coordenada) de P , b es la componente y (o segunda coordenada) y c es la componente z (o tercera coordenada). El punto P de coordenadas $x = a, y = b, z = c$ se designa por $P(a, b, c)$.

El punto sobre el eje x de abscisa a tiene como coordenadas espaciales $(a, 0, 0)$; el punto sobre el eje y de ordenada b tiene como coordenadas espaciales $(0, b, 0)$ y el punto sobre el eje z de cota c tiene como coordenadas espaciales $(0, 0, c)$. Además, dado un punto $P(a, b, c)$, tenemos que el plano perpendicular al eje x que pasa por P corta dicho eje en el punto $(a, 0, 0)$; el plano perpendicular al eje y que pasa por P corta dicho eje en el punto $(0, b, 0)$; el plano perpendicular al eje z que pasa por P corta dicho eje en el punto $(0, 0, c)$.



Un punto está sobre el plano xy si y sólo si es de la forma $(x, y, 0)$. Luego una ecuación para el plano xy es $z = 0$. Claramente, la ecuación $z = z_0$ representa un plano paralelo al plano xy , z_0 unidades por encima (si z_0 es positivo), o por debajo (si z_0 es negativo), del plano xy . De manera análoga, las ecuaciones de la forma $x = x_0$ representan planos paralelos al plano yz y las ecuaciones de la forma $y = y_0$ representan planos paralelos al plano xz .

Por todo lo dicho, podemos identificar la recta con el conjunto de todos los números reales \mathbb{R} ; el plano con el conjunto de todos los pares ordenados (x, y) de números reales \mathbb{R}^2 y, finalmente, el espacio tridimensional con el conjunto de todas las ternas ordenadas (x, y, z) de números reales \mathbb{R}^3 . Se puede hablar de todos estos espacios como \mathbb{R}^n , $n = 1, 2, 3$.

Veamos ahora como podemos extender la operación suma de los números reales de \mathbb{R} , a los pares y las ternas ordenadas de \mathbb{R}^2 y \mathbb{R}^3 . Para \mathbb{R}^3 se procede de la siguiente manera (para \mathbb{R}^2 el procedimiento es completamente análogo).

Definición. Dadas dos ternas $(x, y, z) \in \mathbb{R}^3$ y $(x', y', z') \in \mathbb{R}^3$ definimos su **suma** como la terna de \mathbb{R}^3 tal que:

$$(x, y, z) + (x', y', z') = (x + x', y + y', z + z').$$

El elemento $(0, 0, 0)$ se llama elemento **ceró** de \mathbb{R}^3 . El elemento $(-x, -y, -z)$ se llama **opuesto** de (x, y, z) .

Por ejemplo, dadas las ternas $(2, 1, -3)$ y $(-2, 2, -2)$ tendremos

$$(2, 1, -3) + (-2, 2, -2) = (2 - 2, 1 + 2, -3 - 2) = (0, 3, -5).$$

Hay varias operaciones relacionadas con el producto que son importantes en \mathbb{R}^3 . Dos de ellas se estudiarán con detalle más adelante: el producto interno ó producto escalar, que

9.1. VECTORES EN EL ESPACIO TRIDIMENSIONAL.

asigna un número real a cada pareja de elementos de \mathbb{R}^3 ; y el producto cruz ó producto vectorial, que asigna un elemento de \mathbb{R}^3 a cada pareja de elementos de \mathbb{R}^3 . Otra operación de producto para \mathbb{R}^3 combina escalares (es decir, números reales) con elementos de \mathbb{R}^3 (ternas ordenadas) para producir un elemento de \mathbb{R}^3 .

Definición. Dados un escalar $\alpha \in \mathbb{R}$ y una terna $(x, y, z) \in \mathbb{R}^3$ definimos el **producto por un escalar** como la terna dada por

$$\alpha(x, y, z) = (\alpha x, \alpha y, \alpha z).$$

Por ejemplo, dadas las ternas $(2, 1, -3)$ y $(-2, 2, -2)$ y el escalar $\alpha = 3$ tendremos

$$\alpha(2, 1, -3) = 3(2, 1, -3) = (6, 3, -9), \quad \alpha(-2, 2, -2) = 3(-2, 2, -2) = (-6, 6, -6).$$

Como consecuencia de las definiciones de estas dos operaciones para ternas de \mathbb{R}^3 tenemos las siguientes **propiedades**:

1. Asociatividad:

$$(\alpha\beta)(x, y, z) = \alpha[\beta(x, y, z)].$$

2. Distributividad:

$$\begin{aligned}(\alpha + \beta)(x, y, z) &= \alpha(x, y, z) + \beta(x, y, z); \\ \alpha[(x, y, z) + (x', y', z')] &= \alpha(x, y, z) + \alpha(x', y', z').\end{aligned}$$

3. Propiedades del elemento cero:

$$\alpha(0, 0, 0) = (0, 0, 0) \quad \text{y} \quad 0(x, y, z) = (0, 0, 0).$$

4. Propiedad del elemento identidad:

$$1(x, y, z) = (x, y, z).$$

Vamos a dar ahora el concepto de vector.

Definición. Se define geoméricamente un **vector** como un segmento de recta dirigido que comienza en el origen, es decir, un segmento de recta con magnitud y dirección especificados, con punto inicial en el origen. Los vectores se pueden asociar a flechas que comienzan en el origen. Generalmente se denotan con una letra con flecha encima, \vec{v} o en letras negritas, **v**.

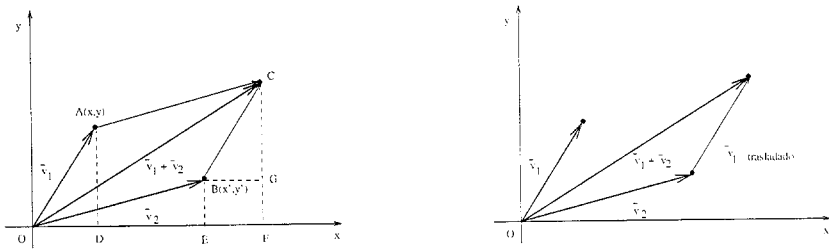
Usando esta definición de vector podemos asociar a cada vector \vec{v} el punto (x, y, z) en el espacio, donde termina \vec{v} , y, recíprocamente, a cada punto (x, y, z) podemos asociar un vector \vec{v} . Es decir, podemos pensar los elementos de \mathbb{R}^3 como ternas ordenadas de números reales o como vectores. La terna $(0, 0, 0)$ es el vector nulo $\vec{0}$.

Definición. Dos vectores $\vec{v}_1 = (x_1, y_1, z_1)$ y $\vec{v}_2 = (x_2, y_2, z_2)$ son iguales si, y sólo si, tienen la misma dirección y la misma magnitud, es decir si coinciden sus componentes:

$$\vec{v}_1 = \vec{v}_2 \quad \iff \quad x_1 = x_2, \quad y_1 = y_2, \quad z_1 = z_2.$$

CAPÍTULO 9. LA GEOMETRÍA DEL ESPACIO EUCLÍDEO.

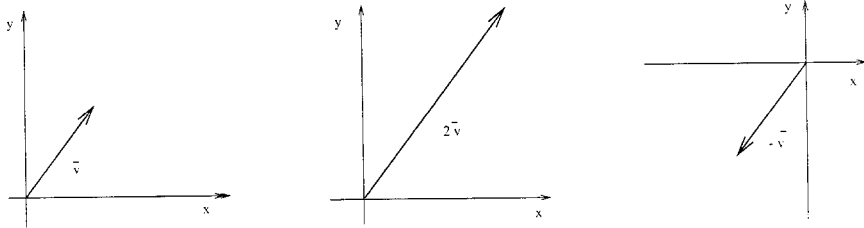
Claramente para los vectores son válidas las definiciones de suma y producto por un escalar dadas anteriormente. La suma de vectores se puede visualizar como la adición de flechas siguiendo la **ley del paralelogramo**: en el plano que contiene los vectores \vec{v}_1 y \vec{v}_2 , formemos el paralelogramo que tiene como un lado a \vec{v}_1 y como lado adyacente a \vec{v}_2 . Entonces, la suma $\vec{v}_1 + \vec{v}_2$ es el segmento de recta dirigido a lo largo de la diagonal del paralelogramo. De manera parecida se puede considerar la suma vectorial en términos de triángulos.



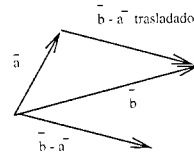
Los múltiplos escalares de los vectores también tienen interpretaciones geométricas. Si α es un escalar y \vec{v} un vector, se define $\alpha \vec{v}$ como el vector que tiene α veces la longitud de \vec{v} , con la misma dirección que \vec{v} si $\alpha > 0$, pero con dirección opuesta si $\alpha < 0$:

$$\alpha \vec{v} = \alpha (x, y, z) = (\alpha x, \alpha y, \alpha z).$$

Claramente los dos vectores \vec{v} y $\alpha \vec{v}$ serán paralelos.



Veamos ahora cual es la representación geométrica de la diferencia de dos vectores \vec{a} y \vec{b} . Como $\vec{a} + (\vec{b} - \vec{a}) = \vec{b}$, resulta que $\vec{b} - \vec{a}$ es el vector que al sumarlo a \vec{a} da \vec{b} . Es decir, será un vector paralelo y con la misma magnitud que el segmento de recta dirigido que comienza en el punto final de \vec{a} y termina en el punto final de \vec{b} .



Si denotemos por \vec{i} al vector que termina en $(1, 0, 0)$, por \vec{j} al vector que termina en $(0, 1, 0)$ y por \vec{k} al vector que termina en $(0, 0, 1)$, tendremos que por la definición de suma vectorial y la multiplicación por un escalar, cualquier vector en el espacio tridimensional

9.1. VECTORES EN EL ESPACIO TRIDIMENSIONAL.

se puede expresar en términos de los vectores \vec{i} , \vec{j} y \vec{k} :

$$\begin{aligned}\vec{v} &= (x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) = \\ &= x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1) = \\ &= x\vec{i} + y\vec{j} + z\vec{k}.\end{aligned}$$

Definición. Los tres vectores \vec{i} , \vec{j} y \vec{k} se llaman vectores coordenados unitarios, o vectores de la base canónica para \mathbb{R}^3 . Los números x , y , z se llaman componentes de \vec{i} , \vec{j} y \vec{k} , respectivamente, del vector \vec{v} .

La suma y la multiplicación por un escalar se pueden escribir en términos de los vectores de la base canónica de la siguiente manera:

$$(x\vec{i} + y\vec{j} + z\vec{k}) + (x'\vec{i} + y'\vec{j} + z'\vec{k}) = (x + x')\vec{i} + (y + y')\vec{j} + (z + z')\vec{k},$$

$$\alpha(x\vec{i} + y\vec{j} + z\vec{k}) = (\alpha x)\vec{i} + (\alpha y)\vec{j} + (\alpha z)\vec{k}.$$

Dados dos vectores \vec{v} , \vec{w} , no paralelos, estos determinan un plano, y además cualquier punto de este plano es de la forma

$$\alpha\vec{v} + \beta\vec{w}, \quad \text{con } \alpha, \beta \in \mathbb{R}, \quad \vec{v}, \vec{w} \in \mathbb{R}^3.$$

El plano determinado por \vec{v} y \vec{w} se llama **plano generado por \vec{v} y \vec{w}** .

Si \vec{v} y \vec{w} son paralelos y ambos no nulos, el plano degenera en una recta.

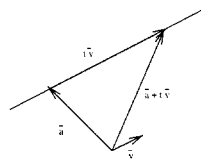
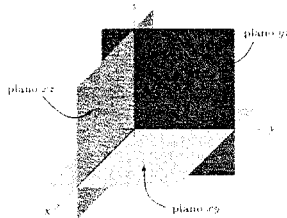
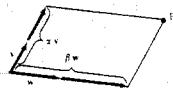
Si $\vec{v} = \vec{w} = \vec{0}$, obtenemos un solo punto.

Hay tres planos particulares que surgen de manera natural en un sistema coordenado y que usaremos muy a menudo: el plano generado por \vec{i} y \vec{j} , que se llama plano xy , el plano generado por \vec{i} y \vec{k} , que se llama plano xz y el plano generado por \vec{j} y \vec{k} , que se llama plano yz . Estos se llaman **planos coordenados**.

Como veremos más adelante, los planos y las rectas son objetos geométricos que se pueden representar mediante ecuaciones. Con el material ahora a disposición somos capaces de hallar **la ecuación de una recta l que pase por el punto final o extremo del vector \vec{a} , con la dirección de un vector \vec{v}** .

Por construcción geométrica vemos que la recta l se puede expresar mediante la ecuación dada en forma vectorial

$$\vec{l}(t) = \vec{a} + t\vec{v}.$$



Decimos que l está expresada de manera paramétrica, con el parámetro t . En $t = 0$, $\vec{l}(0) = \vec{a}$. Cuando t crece, el punto $\vec{l}(t)$ se mueve alejándose de \vec{a} en la dirección de \vec{v} y cuando t decrece desde $t = 0$ hacia los valores negativos, el punto $\vec{l}(t)$ se mueve alejándose de \vec{a} en la dirección de $-\vec{v}$.

Ejemplo. Vamos a escribir la ecuación de la recta que pasa por el punto $\vec{a} = (0, 1, 0)$ con dirección del vector $\vec{v} = \vec{i}$:

$$\vec{l}(t) = \vec{j} + t\vec{i} = (0, 1, 0) + t(1, 0, 0) = (t, 1, 0).$$

Y ahora vamos a deducir la ecuación de la recta que pasa por el punto $(1, -1, 2)$ y es paralela al vector $2\vec{i} - 3\vec{j} + \vec{k}$. Aquí podemos poner $\vec{a} = \vec{i} - \vec{j} + 2\vec{k}$ y $\vec{v} = 2\vec{i} - 3\vec{j} + \vec{k}$. luego

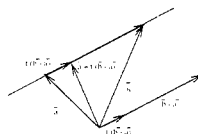
$$\vec{l}(t) = \vec{a} + t\vec{v} = (\vec{i} - \vec{j} + 2\vec{k}) + t(2\vec{i} - 3\vec{j} + \vec{k}) = (1 + 2t)\vec{i} - (1 + 3t)\vec{j} + (2 + t)\vec{k}.$$

Visto que en lugar del punto \vec{a} se puede escoger un punto diferente sobre la recta dada, resulta que puede haber varias parametrizaciones de la misma recta.

Vamos ahora a deducir la ecuación de la **recta que pasa por los puntos finales de dos vectores \vec{a} y \vec{b}** . Como el vector $\vec{b} - \vec{a}$ es paralelo al segmento de recta dirigido que va de \vec{a} a \vec{b} , es lo mismo que calcular la ecuación paramétrica de la recta que pasa por \vec{a} en dirección de $\vec{b} - \vec{a}$:

$$\vec{l}(t) = \vec{a} + t(\vec{b} - \vec{a}) = (1 - t)\vec{a} + t\vec{b}.$$

Conforme t crece de 0 a 1, el vector $\vec{l}(t)$ se mueve de la punta de \vec{a} a la punta de \vec{b} , a lo largo del segmento de recta dirigido de \vec{a} a \vec{b} .



Ejemplo. Escribe la ecuación de la recta que pasa por los puntos $(1, 0, 0)$ y $(1, -1, 2)$.

Solución:

Podemos poner $\vec{a} = (1, 0, 0) = \vec{i}$ y $\vec{b} = (1, -1, 2) = \vec{i} - \vec{j} + 2\vec{k}$. Luego,

$$\vec{l}(t) = (1 - t)\vec{a} + t\vec{b} = (1 - t)\vec{i} + t(\vec{i} - \vec{j} + 2\vec{k}) = \vec{i} - t\vec{j} + 2t\vec{k}.$$

En términos de componentes, la ecuación de la recta que pasa por los puntos (x_1, y_1, z_1) y (x_2, y_2, z_2) es

$$x = x_1 + t(x_2 - x_1), \quad y = y_1 + t(y_2 - y_1), \quad z = z_1 + t(z_2 - z_1).$$

Eliminando t es posible escribir esto como

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

Obviamente, bastan dos cualesquiera de las tres igualdades, y son todas válidas con tal que $x_1 \neq x_2$, $y_1 \neq y_2$, $z_1 \neq z_2$. Si se cumple una de estas igualdades quiere decir que una

9.1. VECTORES EN EL ESPACIO TRIDIMENSIONAL.

de las coordenadas es constante, y desde el punto de vista geométrico esto significa que la recta pertenece a uno de los planos coordenados. Supongamos, por ejemplo, que sea $z_1 = z_2$; entonces las ecuaciones paramétricas escalares toman la forma

$$x = x_1 + t(x_2 - x_1), \quad y = y_1 + t(y_2 - y_1), \quad z = z_1,$$

y eliminando la t , nos quedan dos ecuaciones

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}, \quad z = z_1.$$

La recta está en el plano horizontal $z = z_1$ y su proyección en el plano xy es la recta de ecuación

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} \iff y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1).$$

Hay que hacer notar que cualquier vector de la forma $\vec{c} = \lambda \vec{a} + \mu \vec{b}$, con $\lambda + \mu = 1$, está sobre la recta que pasa por los extremos de \vec{a} y \vec{b} , ya que

$$\vec{c} = \lambda \vec{a} + \mu \vec{b} = (1 - \mu) \vec{a} + \mu \vec{b} = \vec{a} + \mu (\vec{b} - \vec{a}).$$

Ejercicios.

- 1) Dados $\vec{v}_1 = 2\vec{i} - \vec{j} + \vec{k}$, $\vec{v}_2 = \vec{i} + 3\vec{j} - 3\vec{k}$, $\vec{v}_3 = -2\vec{i} + \vec{j} - 3\vec{k}$ y $\vec{v}_4 = 3\vec{i} + 2\vec{j} + 5\vec{k}$, encuentra unos escalares a , b y c tales que $\vec{v}_4 = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3$.
- 2) ¿Qué restricciones se deben tener sobre x , y y z de modo que la terna (x, y, z) represente un punto sobre el eje y ? ¿Y sobre el eje z ? ¿En el plano xz ? ¿En el plano yz ?
- 3) Un ave va volando en línea recta con un vector velocidad $\vec{v} = 10\vec{i} + 6\vec{j} + \vec{k}$ (en kilómetros por hora). Supongamos que (x, y) son sus coordenadas en tierra y z su altura, expresadas en kilómetros.
 - a) Si en un cierto momento el ave está en la posición $(1, 2, 3)$, ¿dónde estará una hora después? ¿Y un minuto después?
 - b) ¿Cuántos segundos tarda el ave en subir 10 metros?
- 4) Encuentra la recta que pasa por $(0, 2, 1)$ en la dirección de $2\vec{i} - \vec{k}$.
- 5) Encuentra la recta que pasa por $(-5, 0, 4)$ y $(6, -3, 2)$.
- 6) Halla los puntos de intersección de la recta $x = 3 + 2t$, $y = 7 + 8t$, $z = -2 + t$, esto es, $\vec{l}(t) = (3 + 2t, 7 + 8t, -2 + t)$, con los planos coordenados.
- 7) Demuestra que si dos vectores \vec{a} y \vec{b} son paralelos a un tercer vector \vec{c} , entonces cualquier combinación lineal de \vec{a} y \vec{b} es también paralela a \vec{c} .

- 8) Sean $\vec{a} = (2, 0, -1)$, $\vec{b} = (1, 3, 5)$, $\vec{c} = (-1, 1, 1)$ y $\vec{d} = (1, 1, 6)$. Expresa $\vec{a} - 3\vec{b} + 2\vec{c} + 4\vec{d}$ como combinación lineal de \vec{i} , \vec{j} , \vec{k} . Halla escalares A , B , C tales que $\vec{d} = A\vec{a} + B\vec{b} + C\vec{c}$.
- 9) Encuentra la recta que es paralela a la recta $\vec{r}(t) = (\vec{i} - \vec{j}) + t\vec{k}$ y pasa por $(3, 1, 0)$. Y una segunda recta que pasa por $(1, -1, 2)$ y es paralela a $\vec{r}(t) = t(3\vec{i} - \vec{j} + \vec{k})$.
- 10) Encuentra el plano generado por $\vec{v}_1 = (2, 7, 0)$ y $\vec{v}_2 = (0, 2, 7)$.
- 11) Muestra que todo punto sobre la recta $\vec{v}(t) = (1, -1, 2) + t(2, 3, 1)$ satisface la ecuación $5x - 3y - z - 6 = 0$.
- 12) Escribe la ecuación química $pC_3H_4O_3 + qO_2 = rCO_2 + sH_2O$ como una ecuación en ternas ordenadas con coeficientes desconocidos p , q , r y s . Halla la menor solución entera positiva posibles para p , q , r y s .

9.2 Producto escalar.

Pasamos a estudiar el primero de los dos productos de vectores que son muy útiles en aplicaciones físicas y tienen interpretaciones geométricas interesantes.

Definición. Sean dos vectores \vec{a} y \vec{b} en \mathbb{R}^3 , con $\vec{a} = (a_1, a_2, a_3) = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$ y $\vec{b} = (b_1, b_2, b_3) = b_1\vec{i} + b_2\vec{j} + b_3\vec{k}$. Definimos el **producto escalar de \vec{a} y \vec{b}** , (producto interno o producto punto), como el número real

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Conviene remarcar que el producto escalar de dos vectores es una cantidad escalar. A partir de la definición se siguen ciertas **propiedades** del producto escalar. Si \vec{a} , \vec{b} y \vec{c} son vectores en \mathbb{R}^3 y α y β son números reales, entonces

1. $\vec{a} \cdot \vec{a} \geq 0$; $\vec{a} \cdot \vec{a} = 0$ si y sólo si $\vec{a} = \vec{0}$;
2. $\alpha \vec{a} \cdot \vec{b} = \alpha (\vec{a} \cdot \vec{b})$ y $\vec{a} \cdot \beta \vec{b} = \beta (\vec{a} \cdot \vec{b})$;
3. $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$ y $(\vec{a} + \vec{b}) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}$;
4. $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$.

La demostración de estas propiedades es muy sencilla y la dejamos como ejercicio. Usando el teorema de Pitágoras, se sigue que la longitud del vector $\vec{a} = (a_1, a_2, a_3)$ es $\sqrt{a_1^2 + a_2^2 + a_3^2}$. Luego podemos dar la siguiente

Definición. La **norma** de un vector $\vec{v} = (v_1, v_2, v_3)$ es el número

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2} = (\vec{v} \cdot \vec{v})^{1/2}.$$

9.2. PRODUCTO ESCALAR.

A la norma de \vec{v} también se le llama longitud o magnitud de \vec{v} . Los vectores que tienen norma 1 se llaman **vectores unitarios**. Por ejemplo, los ya usados vectores \vec{i} , \vec{j} y \vec{k} son vectores unitarios. Observad que para cualquier vector no nulo \vec{v} , el vector $\vec{v}/\|\vec{v}\|$, obtenido dividiendo el vector dado por su norma, es un vector unitario. Se dice que se ha normalizado el vector \vec{v} .

Ejemplo. Sea dado el vector $\vec{v} = (1, 1, 1)$, $\|\vec{v}\| = \sqrt{1+1+1} = \sqrt{3}$. Sea ahora

$$\vec{v}_1 = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\|\vec{v}\|} \vec{v} = \frac{1}{\sqrt{3}} (1, 1, 1) = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) \implies$$

$$\|\vec{v}_1\| = \sqrt{\left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2 + \left(\frac{1}{\sqrt{3}}\right)^2} = \sqrt{\frac{1}{3} + \frac{1}{3} + \frac{1}{3}} = \sqrt{\frac{3}{3}} = 1.$$

Por la construcción del vector $\vec{b} - \vec{a}$, resulta que la distancia del extremo de \vec{a} al extremo de \vec{b} es $\|\vec{b} - \vec{a}\|$.

Ejemplo. Halla la distancia del extremo del vector $\vec{a} = (1, 1, 1)$ al extremo del vector $\vec{b} = (0, 2, 3)$.

Solución:

$$\|\vec{b} - \vec{a}\| = \sqrt{(0-1)^2 + (2-1)^2 + (3-1)^2} = \sqrt{1+1+4} = \sqrt{6}.$$

Con el teorema que sigue vamos a dar la interpretación geométrica del producto escalar: mostraremos que el producto escalar mide el ángulo entre los dos vectores.

Teorema 9.1 Sean \vec{a} y \vec{b} dos vectores en \mathbb{R}^3 y sea θ , con $0 \leq \theta \leq \pi$, el ángulo entre ellos. Entonces

$$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos \theta.$$

De modo que podemos expresar el ángulo entre \vec{a} y \vec{b} como

$$\theta = \arccos \left(\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|} \right).$$

Demostración:

Si aplicamos la ley de los cosenos al triángulo con un vértice en el origen y lados adyacentes determinados por los vectores \vec{a} y \vec{b} , se sigue que

$$\|\vec{b} - \vec{a}\|^2 = \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2 \|\vec{a}\| \|\vec{b}\| \cos \theta$$

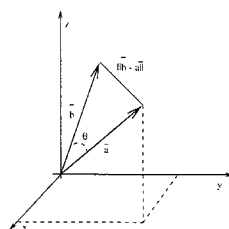
$$(\vec{b} - \vec{a}) \cdot (\vec{b} - \vec{a}) = \vec{a} \cdot \vec{a} + \vec{b} \cdot \vec{b} - 2 \|\vec{a}\| \|\vec{b}\| \cos \theta.$$

El término a la izquierda nos da

$$(\vec{b} - \vec{a}) \cdot (\vec{b} - \vec{a}) = \vec{a} \cdot \vec{a} + \vec{b} \cdot \vec{b} - 2 \vec{a} \cdot \vec{b},$$

de manera que podemos concluir

$$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos \theta.$$



c.q.d.

Corolario 9.2 (Desigualdad de Cauchy-Schwartz).

Para dos vectores \vec{a} y \vec{b} cualesquiera, tenemos

$$|\vec{a} \cdot \vec{b}| \leq \|\vec{a}\| \|\vec{b}\|$$

con la igualdad si y sólo si \vec{a} es un múltiplo escalar de \vec{b} , o uno de ellos es $\vec{0}$.

Demostración:

Si \vec{a} no es un múltiplo escalar de \vec{b} , entonces $|\cos \theta| < 1$ y se cumple la desigualdad.

Cuando \vec{a} es un múltiplo escalar de \vec{b} , entonces $\theta = 0$ o $\theta = \pi$, es decir $|\cos \theta| = 1$.

c.q.d.

Ejemplo. Halla el ángulo entre los vectores $\vec{i} + 2\vec{j}$ e $\vec{i} - \vec{j} + 3\vec{k}$.

Solución:

Usando el teorema 9.1, tenemos

$$(\vec{i} + 2\vec{j}) \cdot (\vec{i} - \vec{j} + 3\vec{k}) = \|\vec{i} + 2\vec{j}\| \|\vec{i} - \vec{j} + 3\vec{k}\| \cos \theta$$

de manera que

$$1 - 2 + 0 = \sqrt{5} \sqrt{11} \cos \theta$$

de donde

$$\cos \theta = \frac{-1}{\sqrt{55}} \iff \theta = \arccos \left(-\frac{1}{\sqrt{55}} \right) \approx 1.71 \text{ radianes } (97.75^\circ).$$

Si \vec{a} y \vec{b} son vectores no nulos de \mathbb{R}^3 y θ el ángulo entre ellos, vemos que $\vec{a} \cdot \vec{b} = 0$ si y sólo si $\cos \theta = 0$. Así, el producto escalar de dos vectores distintos de cero es cero si y sólo si los vectores son perpendiculares, u ortogonales. Los vectores de la base canónica \vec{i} , \vec{j} y \vec{k} son ortogonales entre sí, y tienen norma unitaria, por esto se llaman ortonormales. El vector nulo es ortogonal a todos los vectores.

Ejemplo. Los vectores $\vec{i}_\theta = (\cos \theta) \vec{i} + (\sin \theta) \vec{j}$ y $\vec{j}_\theta = -(\sin \theta) \vec{i} + (\cos \theta) \vec{j}$ son ortogonales, dado que

$$\vec{i}_\theta \cdot \vec{j}_\theta = -\cos \theta \sin \theta + \sin \theta \cos \theta = 0.$$

Vamos ahora a deducir las expresiones para las **proyecciones** de un vector cualquiera sobre dos vectores ortogonales. Dados dos vectores \vec{a} y \vec{b} ortogonales no nulos, si \vec{c} es un vector en el plano generado por \vec{a} y \vec{b} , entonces existen dos escalares α y β tales que $\vec{c} = \alpha \vec{a} + \beta \vec{b}$. Tomando el producto escalar de \vec{a} y \vec{c} , tenemos

$$\vec{a} \cdot \vec{c} = \vec{a} \cdot (\alpha \vec{a} + \beta \vec{b}) = \alpha \vec{a} \cdot \vec{a} + \beta \vec{a} \cdot \vec{b}$$

y como \vec{a} y \vec{b} son ortogonales, $\vec{a} \cdot \vec{b} = 0$, luego

$$\alpha = \frac{\vec{a} \cdot \vec{c}}{\|\vec{a}\|^2}.$$

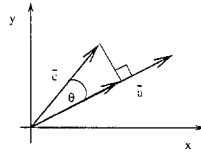
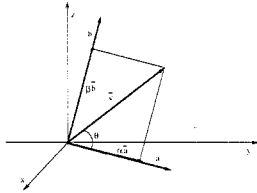
9.2. PRODUCTO ESCALAR.

De la misma manera se obtiene

$$\beta = \frac{\vec{b} \cdot \vec{c}}{\|\vec{b}\|^2}.$$

El vector $\alpha \vec{a}$ se llama la **proyección** de \vec{c} a lo largo de \vec{a} , y $\beta \vec{b}$ se llama la **proyección** de \vec{c} a lo largo de \vec{b} . Notar que la longitud de la proyección de un vector \vec{c} sobre un vector \vec{a} , donde θ es el ángulo entre \vec{a} y \vec{c} , está dada por

$$\|\vec{c}\| \cos \theta = \frac{|\vec{a} \cdot \vec{c}|}{\|\vec{a}\|}.$$



Ejercicios.

- Calcula dos vectores unitarios, no paralelos, ortogonales al vector de componentes $(1,1,1)$.
- Calcula el ángulo entre los dos vectores \vec{v}_1 y \vec{v}_2 si
 - $\vec{v}_1 = 7\vec{j} + 19\vec{k}$ y $\vec{v}_2 = -2\vec{i} - \vec{j}$;
 - $\vec{v}_1 = 3\vec{i} - \vec{j} - 2\vec{k}$ y $\vec{v}_2 = \vec{i} - 2\vec{j} - \vec{k}$;
 - $\vec{v}_1 = 2\vec{i} - 3\vec{j} + \vec{k}$ y $\vec{v}_2 = -3\vec{i} + \vec{j} + 9\vec{k}$.
- Determina un vector unitario perpendicular al plano que contenga $\vec{v}_1 = 2\vec{i} - 6\vec{j} - 3\vec{k}$ y $\vec{v}_2 = 4\vec{i} + 3\vec{j} - \vec{k}$.
- Encuentra el trabajo hecho sobre un objeto que se mueve a lo largo del vector $\vec{v} = 3\vec{i} + 2\vec{j} - 5\vec{k}$, si la fuerza aplicada es $\vec{F} = 2\vec{i} - \vec{j} - \vec{k}$.
- Encuentra la ecuación del plano perpendicular al vector $\vec{A} = 2\vec{i} + 3\vec{j} + 6\vec{k}$ y que pasa por el extremo final del vector $\vec{B} = \vec{i} + 5\vec{j} + 3\vec{k}$.
- Halla α sabiendo que
 - $\vec{v}_1 = 3\vec{i} + \vec{j} - \vec{k}$ y $\vec{v}_2 = \alpha\vec{i} - 4\vec{j} + 4\vec{k}$ son paralelos;
 - $\vec{v}_1 = 3\vec{i} + \vec{j}$ y $\vec{v}_2 = \alpha\vec{j} - \vec{k}$ tienen la misma longitud;
 - $\|\alpha\vec{i} + (\alpha - 1)\vec{j} + (\alpha + 1)\vec{k}\| = 2$;
 - $\vec{v}_1 = \alpha\vec{i} + 11\vec{j} - 3\vec{k}$ y $\vec{v}_2 = 2\alpha\vec{i} - \alpha\vec{j} - 5\vec{k}$ son perpendiculares.

- 7) Dados los vectores $\vec{a} = 2\vec{i} + \vec{j}$, $\vec{b} = 3\vec{i} - \vec{j} + 2\vec{k}$ y $\vec{c} = 4\vec{i} + 3\vec{k}$. calcula
- los tres productos escalares $\vec{a} \cdot \vec{b}$, $\vec{a} \cdot \vec{c}$, $\vec{b} \cdot \vec{c}$;
 - los cosenos de los ángulos formados por estos vectores;
 - las componentes de \vec{a} en las direcciones de \vec{b} y \vec{c} ;
 - las proyecciones de \vec{a} en las direcciones de \vec{b} y \vec{c} .
- 8) Halla los vectores unitarios \vec{u} que son perpendiculares a la vez a $\vec{v}_1 = \vec{i} + 2\vec{j} + \vec{k}$ y $\vec{v}_2 = 3\vec{i} - 4\vec{j} + 2\vec{k}$.

9.3 Producto vectorial.

En la sección anterior hemos definido un producto de vectores que daba como resultado un escalar. Ahora vamos a definir un producto de vectores cuyo resultado es otro vector.

Definición. Sean \vec{a} y \vec{b} dos vectores en \mathbb{R}^3 , con $\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$ y $\vec{b} = b_1\vec{i} + b_2\vec{j} + b_3\vec{k}$. El **producto vectorial de \vec{a} y \vec{b}** , denotado por $\vec{a} \times \vec{b}$, está definido como el vector

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \vec{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \vec{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \vec{k}.$$

Como veremos este nuevo vector tendrá la propiedad geométrica de ser perpendicular al plano generado por \vec{a} y \vec{b} . Pero antes veamos ciertas **propiedades** algebraicas del producto vectorial que se deducen de su definición. Si \vec{a} , \vec{b} y \vec{c} son vectores y α , β y γ son escalares, entonces

1. Anticonmutativa:

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}.$$

2. Factorización de los escalares:

$$\alpha \vec{a} \times \beta \vec{b} = \alpha \beta (\vec{a} \times \vec{b}).$$

3. Distributividad:

$$\vec{a} \times (\beta \vec{b} + \gamma \vec{c}) = \beta (\vec{a} \times \vec{b}) + \gamma (\vec{a} \times \vec{c});$$

$$(\alpha \vec{a} + \beta \vec{b}) \times \vec{c} = \alpha (\vec{a} \times \vec{c}) + \beta (\vec{b} \times \vec{c}).$$

Podemos notar que por la propiedad 1. $\vec{a} \times \vec{a} = -\vec{a} \times \vec{a}$, es decir, $\vec{a} \times \vec{a} = \vec{0}$. En particular, tenemos

$$\vec{i} \times \vec{i} = \vec{0}, \quad \vec{j} \times \vec{j} = \vec{0}, \quad \vec{k} \times \vec{k} = \vec{0}.$$

9.3. PRODUCTO VECTORIAL.

Además

$$\vec{i} \times \vec{j} = \vec{k}, \quad \vec{j} \times \vec{k} = \vec{i}, \quad \vec{k} \times \vec{i} = \vec{j}.$$

Queremos ahora dar una interpretación geométrica del producto vectorial. Para hacerlo, introducimos antes el triple producto o producto mixto.

Definición. Dados tres vectores \vec{a} , \vec{b} y \vec{c} en \mathbb{R}^3 , al número real

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} a_1 - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} a_2 + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} a_3,$$

se llama **producto triple de \vec{a} , \vec{b} y \vec{c}** .

Si el vector \vec{a} es un vector en el plano generado por los vectores \vec{b} y \vec{c} , quiere decir que podemos escribir $\vec{a} = \alpha \vec{b} + \beta \vec{c}$ y por lo visto en la sección 8.5 si en un determinante una de sus líneas es combinación lineal de otras paralelas a ella, el determinante es nulo. Por lo tanto, $\vec{a} \cdot (\vec{b} \times \vec{c}) = 0$; en otras palabras el vector $\vec{b} \times \vec{c}$ es ortogonal a cualquier vector en el plano generado por \vec{b} y \vec{c} , en particular es ortogonal tanto a \vec{b} como a \vec{c} .

Veamos ahora la magnitud del vector $\vec{b} \times \vec{c}$:

$$\begin{aligned} \|\vec{b} \times \vec{c}\|^2 &= \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}^2 = \\ &= (b_2 c_3 - b_3 c_2)^2 + (b_1 c_3 - b_3 c_1)^2 + (b_1 c_2 - b_2 c_1)^2 = \\ &= (b_1^2 + b_2^2 + b_3^2)(c_1^2 + c_2^2 + c_3^2) - (b_1 c_1 + b_2 c_2 + b_3 c_3)^2 = \\ &= \|\vec{b}\|^2 \|\vec{c}\|^2 - (\vec{b} \cdot \vec{c})^2 = \|\vec{b}\|^2 \|\vec{c}\|^2 - \|\vec{b}\|^2 \|\vec{c}\|^2 \cos^2 \theta = \\ &= \|\vec{b}\|^2 \|\vec{c}\|^2 \sin^2 \theta, \end{aligned}$$

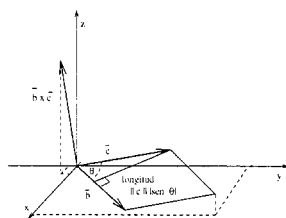
donde θ es el ángulo entre \vec{b} y \vec{c} , con $0 \leq \theta \leq \pi$.

Podemos entonces concluir que $\vec{b} \times \vec{c}$ es un vector perpendicular al plano generado por \vec{b} y \vec{c} , con longitud $\|\vec{b}\| \|\vec{c}\| \sin \theta$.

Sin embargo, hay dos vectores que pueden satisfacer estas condiciones, \vec{n}_1 y $\vec{n}_2 = -\vec{n}_1$ con $\|\vec{n}_1\| = \|\vec{n}_2\| = \|\vec{b}\| \|\vec{c}\| \sin \theta$.

Se elige entre los dos el que sigue la regla de la mano derecha.

Si \vec{b} y \vec{c} son colineales, $\sin \theta = 0$, de modo que $\vec{b} \times \vec{c} = \vec{0}$. Si \vec{b} y \vec{c} no son colineales, entonces generan un plano, de modo que $\vec{b} \times \vec{c}$ es un vector perpendicular a este plano; su longitud $\|\vec{b}\| \|\vec{c}\| \sin \theta$, es simplemente el área del paralelogramo que tiene como lados adyacentes a los vectores \vec{b} y \vec{c} .



Resulta además que dados tres vectores \vec{a} , \vec{b} y \vec{c} en \mathbb{R}^3 el valor absoluto del producto mixto $\vec{a} \cdot (\vec{b} \times \vec{c})$ es el volumen del paralelepípedo con aristas adyacentes \vec{a} , \vec{b} y \vec{c} . Sabemos

que $\|\vec{b} \times \vec{c}\|$ es el área del paralelogramo con lados adyacentes \vec{b} y \vec{c} . Además sabemos que $|\vec{a} \cdot (\vec{b} \times \vec{c})| = \|\vec{a}\| \|\vec{b} \times \vec{c}\| \cos \psi$, donde ψ es el ángulo agudo que forma \vec{a} con la normal al plano generado por \vec{b} y \vec{c} . Y esta última expresión resulta ser el volumen del paralelepípedo: producto del área de la base $\|\vec{b} \times \vec{c}\|$ por la altura $\|\vec{a}\| \cos \psi$.

Ejemplo 1. Halla un vector unitario ortogonal a los vectores $\vec{i} + \vec{k}$ y $\vec{j} + \vec{k}$.

Solución:

Por lo visto arriba, un vector ortogonal a los dos dados será

$$(\vec{i} + \vec{k}) \times (\vec{j} + \vec{k}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = -\vec{i} - \vec{j} + \vec{k}.$$

Como $\|-\vec{i} - \vec{j} + \vec{k}\| = \sqrt{3}$, el vector unitario perpendicular a $\vec{i} + \vec{k}$ y $\vec{j} + \vec{k}$ será

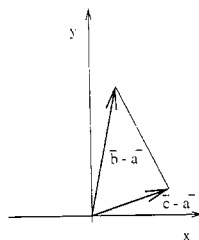
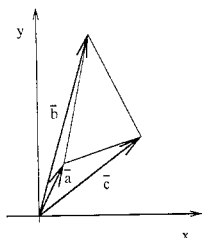
$$-\frac{1}{\sqrt{3}} (\vec{i} + \vec{j} - \vec{k}).$$

Ejemplo 2. Halla el área del triángulo con vértices en los puntos (1, 2), (2, 7) y (4, 3).

Solución:

Sean $\vec{a} = \vec{i} + 2\vec{j}$, $\vec{b} = 2\vec{i} + 7\vec{j}$ y $\vec{c} = 4\vec{i} + 3\vec{j}$.

Por traslación el triángulo de vértices los extremos de los vectores \vec{a} , \vec{b} y \vec{c} tiene la misma área que el triángulo de vértices en $\vec{0}$, $\vec{b} - \vec{a}$ y $\vec{c} - \vec{a}$.



Como el área de este triángulo trasladado es la mitad del área del paralelogramo con lados adyacentes $\vec{v}_1 = \vec{b} - \vec{a} = \vec{i} + 5\vec{j}$ y $\vec{v}_2 = \vec{c} - \vec{a} = 3\vec{i} + \vec{j}$, tenemos que el área buscada es la mitad de la magnitud del producto vectorial entre estos dos vectores:

$$\text{Área} = \frac{1}{2} \|\vec{v}_1 \times \vec{v}_2\|.$$

Ahora tenemos que

$$\vec{v}_1 \times \vec{v}_2 = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 5 & 0 \\ 3 & 1 & 0 \end{vmatrix} = -14\vec{k} \implies \|\vec{v}_1 \times \vec{v}_2\| = 14 \implies \text{Área} = \frac{14}{2} = 7.$$

9.3. PRODUCTO VECTORIAL.

Ejemplo 3. Halla el volumen del paralelepípedo con lados $\vec{a} = \vec{i} + 2\vec{j} + 3\vec{k}$, $\vec{b} = \vec{i} + 3\vec{j}$ y $\vec{c} = 2\vec{i} + \vec{j}$.

Solución:

El producto mixto vale

$$\vec{a} \cdot (\vec{b} \times \vec{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 3 & 0 \\ 2 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} 3 = 3(1 - 6) = -15;$$

luego, su valor absoluto nos da el volumen del paralelepípedo: $V = 15$.

Para concluir esta sección usaremos ahora métodos vectoriales para determinar la ecuación de un plano en el espacio.

Sean P un plano en el espacio, \vec{a} un vector que termina en el plano, y \vec{n} un vector normal al plano. Si \vec{r} es un vector de \mathbb{R}^3 , entonces el extremo de \vec{r} está en el plano si, y sólo si, el vector $\vec{r} - \vec{a}$ es paralelo al plano P , y por lo tanto si, y sólo si,

$$(\vec{r} - \vec{a}) \cdot \vec{n} = 0.$$

Esta última condición es equivalente a $\vec{r} \cdot \vec{n} = \vec{a} \cdot \vec{n}$.

Si ponemos $\vec{a} = a_1\vec{i} + a_2\vec{j} + a_3\vec{k}$, $\vec{n} = A\vec{i} + B\vec{j} + C\vec{k}$ y $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$, se sigue que el extremo de \vec{r} está en P si, y sólo si,

$$Ax + By + Cz = \vec{r} \cdot \vec{n} = \vec{a} \cdot \vec{n} = Aa_1 + Ba_2 + Ca_3 = -D$$

donde D es la constante $D = -Aa_1 - Ba_2 - Ca_3$. Entonces, una ecuación que determina el plano P es

$$Ax + By + Cz + D = 0,$$

donde $\vec{n} = A\vec{i} + B\vec{j} + C\vec{k}$ es un vector normal a P ; recíprocamente si A , B y C no son cero simultáneamente, el conjunto de puntos (x, y, z) que satisface la ecuación $Ax + By + Cz + D = 0$ es un plano con normal $\vec{n} = A\vec{i} + B\vec{j} + C\vec{k}$. Los cuatro números A , B , C y D no están determinados de manera única por el plano P : están determinados salvo por un múltiplo escalar, dado que dados A , B , C y D , y A' , B' , C' y D' , determinan el mismo plano si, y sólo si, $A = \alpha A'$, $B = \alpha B'$, $C = \alpha C'$ y $D = \alpha D'$ para algún escalar α .

Finalmente el plano con normal $\vec{n} = A\vec{i} + B\vec{j} + C\vec{k}$ que pasa por un punto $R = (x_0, y_0, z_0)$ es

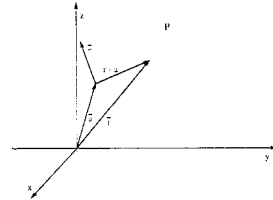
$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Ejemplo 1. Determina la ecuación del plano perpendicular al vector $\vec{i} + 3\vec{j} + 2\vec{k}$, que contiene al punto $(1, 1, 1)$.

Solución:

Usando la última ecuación dada tenemos

$$(x - 1) + 3(y - 1) + 2(z - 1) = 0 \quad \iff \quad x + 3y + 2z = 6.$$



Ejemplo 2. Halla la ecuación del plano que contiene a los puntos $(1, 1, 1)$, $(1, 3, -1)$ y $(2, 1, -2)$.

Solución:

Método 1. Cualquier ecuación del plano es de la forma $Ax + By + Cz + D = 0$. Como los puntos dados están en el plano, sustituyendo en la ecuación del plano tenemos el sistema

$$\begin{cases} A + B + C + D = 0 \\ A + 3B - C + D = 0 \\ 2A + B - 2C + D = 0 \end{cases} \iff \begin{cases} A = 3C \\ B = C \\ D = -5C \end{cases}$$

Como los números A , B , C y D están determinados salvo por un múltiplo escalar, podemos fijar el valor de uno y así los otros quedarán determinados de manera única. Si fijamos $C = 1$, tenemos $A = 3$, $B = 1$ y $D = -5$, es decir la ecuación del plano será

$$3x + y + z = 5.$$

Método 2. Sean \vec{a} , \vec{b} y \vec{c} los tres vectores con extremos en los puntos dados:

$$\vec{a} = \vec{i} + \vec{j} + \vec{k}, \quad \vec{b} = \vec{i} + 3\vec{j} - \vec{k}, \quad \vec{c} = 2\vec{i} + \vec{j} - 2\vec{k}.$$

Cualquier vector normal al plano debe ser ortogonal a los vectores $\vec{a} - \vec{b}$ y $\vec{c} - \vec{b}$, que son paralelos al plano. Así, $\vec{n} = (\vec{a} - \vec{b}) \times (\vec{c} - \vec{b})$ es normal al plano. Calculando, tenemos

$$\vec{a} - \vec{b} = -2\vec{j} + 2\vec{k}, \quad \vec{c} - \vec{b} = \vec{i} - 2\vec{j} - \vec{k},$$

luego

$$(\vec{a} - \vec{b}) \times (\vec{c} - \vec{b}) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 & -2 & 2 \\ 1 & -2 & -1 \end{vmatrix} = 6\vec{i} + 2\vec{j} + 2\vec{k}.$$

Entonces, cualquier ecuación del plano es de la forma $6x + 2y + 2z + D = 0$. Como el punto $(1, 1, 1)$ está en el plano podemos determinar el valor de D :

$$6 + 2 + 2 + D = 0 \quad \implies \quad D = -10$$

entonces la ecuación del plano será

$$6x + 2y + 2z = 10 \quad \iff \quad 3x + y + z = 5.$$

Ejemplo 3. Determina la distancia del punto genérico $E(x_1, y_1, z_1)$ al plano con ecuación $A(x - x_0) + B(y - y_0) + C(z - z_0) = Ax + By + Cz + D = 0$.

Solución:

El vector

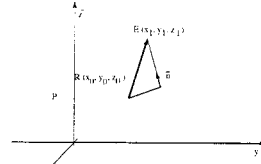
$$\vec{n} = \frac{A\vec{i} + B\vec{j} + C\vec{k}}{\sqrt{A^2 + B^2 + C^2}};$$

9.3. PRODUCTO VECTORIAL.

es un vector unitario normal al plano.

Si bajamos una perpendicular de E al plano y construimos el triángulo REQ , resulta que la distancia $d = |EQ|$ es la longitud de la proyección de $\vec{v} = \vec{RE}$ sobre \vec{n} ; así,

$$d = |\vec{v} \cdot \vec{n}| = \left| \left[(x_1 - x_0) \vec{i} + (y_1 - y_0) \vec{j} + (z_1 - z_0) \vec{k} \right] \cdot \vec{n} \right| = \frac{|A(x_1 - x_0) + B(y_1 - y_0) + C(z_1 - z_0)|}{\sqrt{A^2 + B^2 + C^2}} = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}$$



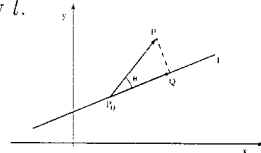
Ejemplo 4. Determina la distancia a una recta l del punto genérico P que no está en l .

Solución:

Estamos interesados en hallar la distancia $d(P, l)$ entre P y l .

Con P_0 y Q como en la figura, podemos tomar $\vec{d} = \vec{P_0Q}$ como vector director de l . Entonces, a partir de la figura se puede ver que

$$d(P, l) = \|\vec{QP}\| = \|\vec{P_0P}\| \operatorname{sen} \theta.$$



Dado que

$$\|\vec{P_0P} \times \vec{d}\| = \|\vec{P_0P}\| \|\vec{d}\| \operatorname{sen} \theta$$

resulta

$$d(P, l) = \frac{\|\vec{P_0P} \times \vec{d}\|}{\|\vec{d}\|} = \frac{\|\vec{P_0P} \times \vec{P_0Q}\|}{\|\vec{P_0Q}\|}.$$

Esta pequeña y elegante fórmula da la distancia de un punto P a cualquier recta l en función de cualquier punto P_0 en l y de cualquier vector director \vec{d} de l .

Ejercicios.

- 1) Halla el ángulo que forman los planos $2x + 3y - z = 2$ y $x - y + z = 1$.
- 2) Calcula $\vec{a} \cdot (\vec{b} \times \vec{c})$ siendo $\vec{a} = \vec{i} - 2\vec{j} + \vec{k}$, $\vec{b} = 2\vec{i} + \vec{j} + \vec{k}$ y $\vec{c} = 3\vec{i} - \vec{j} + 2\vec{k}$.
- 3) Encuentra la ecuación del plano que pasa por $(1, 1, 1)$, $(2, 0, -1)$ y $(0, 4, -3)$.
- 4) Determina un vector unitario perpendicular al plano que contenga $\vec{v}_1 = 2\vec{i} - 6\vec{j} - 3\vec{k}$ y $\vec{v}_2 = 4\vec{i} + 3\vec{j} - \vec{k}$.
- 5) Prueba que
 - a) $\vec{A} \cdot (\vec{B} \times \vec{C}) = \vec{B} \cdot (\vec{C} \times \vec{A}) = \vec{C} \cdot (\vec{A} \times \vec{B})$;
 - b) $\vec{A} \cdot (\vec{B} \times \vec{C}) = (\vec{A} \times \vec{B}) \cdot \vec{C}$;
 - c) $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B} (\vec{A} \cdot \vec{C}) - \vec{C} (\vec{A} \cdot \vec{B})$;

- d) $(\vec{A} \times \vec{B}) \times \vec{C} = \vec{B} (\vec{A} \cdot \vec{C}) - \vec{A} (\vec{B} \cdot \vec{C})$;
- e) $(\vec{A} \times \vec{B}) \cdot (\vec{C} \times \vec{D}) = (\vec{A} \cdot \vec{C}) (\vec{B} \cdot \vec{D}) - (\vec{A} \cdot \vec{D}) (\vec{B} \cdot \vec{C})$;
- f) $\vec{A} \times (\vec{B} \times \vec{C}) + \vec{B} \times (\vec{C} \times \vec{A}) + \vec{C} \times (\vec{A} \times \vec{B}) = \vec{0}$;
- g) $(\vec{A} \times \vec{B}) \times (\vec{C} \times \vec{D}) = \vec{B} (\vec{A} \cdot \vec{C} \times \vec{D}) - \vec{A} (\vec{B} \cdot \vec{C} \times \vec{D}) =$
 $= \vec{C} (\vec{A} \cdot \vec{B} \times \vec{D}) - \vec{D} (\vec{A} \cdot \vec{B} \times \vec{C})$;
- h) $(\vec{A} \times \vec{B}) \cdot (\vec{B} \times \vec{C}) \times (\vec{C} \times \vec{A}) = (\vec{A} \cdot \vec{B} \times \vec{C})^2$.
- 6) Halla los vectores unitarios \vec{u} que son perpendiculares a la vez a \vec{v}_1 y \vec{v}_2 si
- a) $\vec{v}_1 = -5\vec{i} + 9\vec{j} - 4\vec{k}$ y $\vec{v}_2 = 7\vec{i} + 8\vec{j} + 9\vec{k}$;
- b) $\vec{v}_1 = 2\vec{i} - 4\vec{j} + 3\vec{k}$ y $\vec{v}_2 = -4\vec{i} + 8\vec{j} - 6\vec{k}$.
- 7) Calcula $\vec{u} + \vec{v}$, $\vec{u} \cdot \vec{v}$, $\|\vec{u}\|$, $\|\vec{v}\|$ y $\vec{u} \times \vec{v}$, donde $\vec{u} = \vec{i} - 2\vec{j} + \vec{k}$ y $\vec{v} = 2\vec{i} - \vec{j} + 2\vec{k}$.
- 8) Halla una ecuación para el plano que pase por
- a) $(1, 1, 1)$ y sea perpendicular a $\vec{v} = (1, 2, 3)$;
- b) $(5, -1, 0)$ y sea perpendicular a la recta $\vec{l}(t) = (5, 0, 2)t + (3, -1, 1)$;
- c) $(1, 2, -3)$ y sea perpendicular a la recta $\vec{l}(t) = (0, -2, 1) + (1, -2, 3)t$;
- d) $(1, -2, -3)$ y sea perpendicular al plano $3x - y - 2z + 4 = 0$;
- e) $(2, -2, 3)$ y sea perpendicular al plano xz ;
- f) $(1, 4, -3)$ y sea perpendicular al plano yz ;
- g) $(3, 2, -1)$ y $(1, -1, 2)$ y sea paralelo a la recta $\vec{l}(t) = (3, 2, -2)t + (1, -1, 0)$;
- h) $(2, 3, 4)$ y sea perpendicular al vector $\vec{i} - 4\vec{j} + 3\vec{k}$;
- i) $(2, 1, 1)$ y sea paralelo al plano $3x - 2y + 5z - 2 = 0$.
- 9) Halla el área del triángulo PQR, donde $P(0, 1, 0)$, $Q(-1, 1, 2)$ y $R(2, 1, -1)$.
- 10) Halla la distancia del punto $(1, 2, 3)$ a la recta $\vec{r}(t) = \vec{i} + 2\vec{k} + t(\vec{i} - 2\vec{j} + 3\vec{k})$ y del punto $(0, 0, 0)$ a la recta $\vec{r}(t) = \vec{i} + t\vec{j}$.
- 11) Sea l la recta determinada por los puntos P_1 y P_2 , y sea p el plano determinado por los puntos Q_1 , Q_2 y Q_3 . Encuentra, si existe, la intersección de l y p , siendo
- a) $P_1(1, -1, 2)$, $P_2(-2, 3, 1)$, $Q_1(2, 0, -4)$, $Q_2(1, 2, 3)$, $Q_3(-1, 2, 1)$;
- b) $P_1(4, -3, 1)$, $P_2(2, -2, 3)$, $Q_1(2, 0, -4)$, $Q_2(1, 2, 3)$, $Q_3(-1, 2, 1)$.

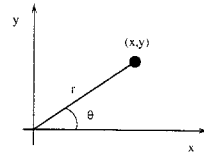
9.4 Coordenadas esféricas y cilíndricas.

Un punto de \mathbb{R}^2 viene representado usualmente usando las coordenadas cartesianas o rectangulares (x, y) . Sin embargo, las coordenadas polares en el plano pueden ser muy útiles.

Definición. Las coordenadas polares (r, θ) están relacionadas con las coordenadas (x, y) mediante las fórmulas

$$x = r \cos \theta, \quad y = r \operatorname{sen} \theta, \quad (9.1)$$

donde usualmente tomamos $r \geq 0$, $0 \leq \theta < 2\pi$. Vamos a hacer algo parecido con los puntos de \mathbb{R}^3 . Además de las coordenadas cartesianas rectangulares (x, y, z) para los puntos del espacio, podemos usar otros sistemas alternativos que serán muy útiles más adelante en el cálculo de las integrales.



Definición. Las coordenadas cilíndricas (r, θ, z) de un punto (x, y, z) vienen definidas por:

$$x = r \cos \theta, \quad y = r \operatorname{sen} \theta, \quad z = z \quad (9.2)$$

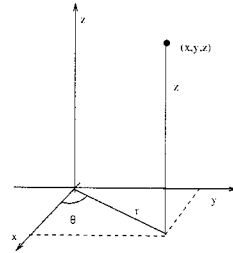
con $r \geq 0$, $0 \leq \theta < 2\pi$ y $-\infty < z < \infty$; o, explícitamente,

$$r = \sqrt{x^2 + y^2}, \quad z = z, \quad \theta = \begin{cases} \operatorname{arctg}(y/x) & \text{si } x > 0 \text{ e } y \geq 0, \\ \pi + \operatorname{arctg}(y/x) & \text{si } x < 0, \\ 2\pi + \operatorname{arctg}(y/x) & \text{si } x > 0 \text{ e } y < 0. \end{cases}$$

Si $x = 0$, entonces $\theta = \pi/2$ para $y > 0$ y $\theta = 3\pi/2$ para $y < 0$. Si $x = y = 0$, θ no está definido.

Las dos primeras coordenadas, r y θ , son las coordenadas polares usuales. La tercera de estas coordenadas es la coordenada rectangular z , que asume todos los valores reales posibles.

En coordenadas rectangulares, las superficies coordenadas $x = x_0$, $y = y_0$, $z = z_0$, son tres planos mutuamente perpendiculares, esto es perpendiculares dos a dos.



En coordenadas cilíndricas, tenemos las tres superficies $r = r_0$, $\theta = \theta_0$, $z = z_0$. La primera, $r = r_0$, representa un cilindro circular de radio r_0 , con el eje z como eje central; la segunda, $\theta = \theta_0$, es un semiplano vertical limitado por el eje z que forma un ángulo de θ_0 radianes con el eje x ; y la última superficie es el plano $z = z_0$ de las coordenadas rectangulares.

Los sólidos más sencillos de describir en términos de las coordenadas cilíndricas son las *cañas cilíndricas*, que consisten en los puntos (x, y, z) cuyas coordenadas cilíndricas (r, θ, z) están en el conjunto $C = \{a_1 \leq r \leq a_2, b_1 \leq \theta \leq b_2, c_1 \leq z \leq c_2\}$.

Ejemplo 1. Halla las coordenadas cilíndricas correspondientes al punto $(3, 3, -2)$.

Solución:

Según la definición de coordenadas cilíndricas, tenemos

$$r = \sqrt{3^2 + 3^2} = 3\sqrt{2} \quad \text{y} \quad \theta = \arctg\left(\frac{3}{3}\right) = \arctg 1 = \frac{\pi}{4},$$

luego las coordenadas cilíndricas del punto $(3, 3, -2)$ son $\left(3\sqrt{2}, \frac{\pi}{4}, -2\right)$.

Ejemplo 2. Si un punto tiene coordenadas cilíndricas $(4, \frac{2}{3}\pi, -3)$ ¿cuáles son sus coordenadas cartesianas?

Solución:

Otra vez, usando la definición tenemos

$$x = r \cos \theta = 4 \cos\left(\frac{2}{3}\pi\right) = 4\left(-\frac{1}{2}\right) = -2$$

y

$$y = r \operatorname{sen} \theta = 4 \operatorname{sen}\left(\frac{2}{3}\pi\right) = 4\left(\frac{\sqrt{3}}{2}\right) = 2\sqrt{3};$$

las coordenadas cartesianas del punto con coordenadas cilíndricas $(4, \frac{2}{3}\pi, -3)$ son

$$(-2, 2\sqrt{3}, -3).$$

Definición. Las **coordenadas esféricas** (ρ, θ, ϕ) de un punto (x, y, z) vienen definidas por:

$$x = \rho \cos \theta \operatorname{sen} \phi, \quad y = \rho \operatorname{sen} \theta \operatorname{sen} \phi, \quad z = \rho \cos \phi \quad (9.3)$$

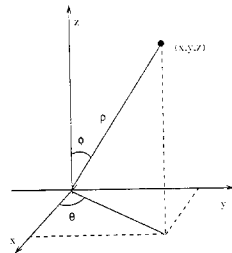
con $\rho \geq 0$, $0 \leq \theta < 2\pi$ y $0 \leq \phi \leq \pi$; o, explícitamente,

$$\rho = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \begin{cases} \arctg(y/x) & \text{si } x > 0 \text{ e } y \geq 0, \\ \pi + \arctg(y/x) & \text{si } x < 0, \\ 2\pi + \arctg(y/x) & \text{si } x > 0 \text{ e } y < 0, \end{cases} \quad \phi = \arccos\left(\frac{z}{\rho}\right).$$

La primera de estas coordenadas, ρ , es la distancia al origen, luego $\rho \geq 0$.

La segunda, θ , es el mismo ángulo de las coordenadas cilíndricas, con $0 \leq \theta < 2\pi$.

La tercera coordenada es el ángulo ϕ , que varía entre $0 \leq \phi \leq \pi$.



9.4. COORDENADAS ESFÉRICAS Y CILÍNDRICAS.

Estas últimas dos coordenadas esféricas, θ y ϕ , se parecen a las coordenadas geográficas de longitud y latitud, si consideramos al eje de la Tierra como el eje z .

Sin embargo, hay diferencias: la longitud geográfica es $|\theta|$ y se llama longitud este u oeste, dependiendo de si θ es positivo o negativo; la latitud geográfica es $|\pi/2 - \phi|$ y se llama latitud norte o sur, dependiendo de si $\pi/2 - \phi$ es positivo o negativo.

En coordenadas esféricas, tenemos las tres superficies $\rho = \rho_0$, $\theta = \theta_0$, $\phi = \phi_0$. La primera, $\rho = \rho_0$, representa una esfera de radio ρ_0 , con centro en el origen; la segunda, $\theta = \theta_0$, es, como en el caso de coordenadas cilíndricas, un semiplano vertical limitado por el eje z que forma un ángulo de θ_0 radianes con el eje x ; y la última superficie, $\phi = \phi_0$, necesita un poco más de explicación. Si $0 < \phi < \pi/2$ o si $\pi/2 < \phi < \pi$, la superficie es la hoja de un cono que se genera al hacer girar alrededor del eje z cualquier radio que nace en el origen y forma ϕ_0 radianes con el eje z . La superficie $\phi = \pi/2$ es el plano xy : la hoja del cono se ha abierto por completo. La ecuación $\phi = 0$ da el semieje de los z positivos, mientras que la ecuación $\phi = \pi$ da el semieje de los z negativos: la hoja del cono se ha cerrado por completo.

Los sólidos más sencillos de describir en términos de las coordenadas esféricas son las *cuñas esféricas*, que consisten en los puntos (x, y, z) cuyas coordenadas esféricas (ρ, θ, ϕ) están en el conjunto $C = \{a_1 \leq \rho \leq a_2, b_1 \leq \theta \leq b_2, c_1 \leq \phi \leq c_2\}$.

Ejemplo 1. Halla las coordenadas esféricas correspondientes al punto $(3, 3, -2)$.

Solución:

Según la definición de coordenadas esféricas, tenemos

$$\rho = \sqrt{3^2 + 3^2 + (-2)^2} = \sqrt{22}, \quad \theta = \arctg \left(\frac{3}{3} \right) = \arctg 1 = \frac{\pi}{4}$$

y

$$\phi = \arccos \left(\frac{-2}{\sqrt{22}} \right) \approx 1.13 \text{ radianes} \approx 64^\circ$$

luego las coordenadas esféricas del punto $(3, 3, -2)$ son $\left(\sqrt{22}, \frac{\pi}{4}, 1.13 \right)$.

Ejemplo 2. Si un punto tiene coordenadas esféricas $\left(4, \frac{2}{3}\pi, \frac{1}{4}\pi \right)$ ¿cuáles son sus coordenadas cartesianas?

Solución:

Otra vez, usando la definición tenemos

$$x = \rho \cos \theta \sin \phi = 4 \cos \left(\frac{2}{3}\pi \right) \sin \left(\frac{1}{4}\pi \right) = 4 \left(-\frac{1}{2} \right) \left(\frac{\sqrt{2}}{2} \right) = -\sqrt{2}$$

$$y = \rho \sin \theta \sin \phi = 4 \sin \left(\frac{2}{3}\pi \right) \sin \left(\frac{1}{4}\pi \right) = 4 \left(\frac{\sqrt{3}}{2} \right) \left(\frac{\sqrt{2}}{2} \right) = \sqrt{6};$$

y

$$z = \rho \cos \phi = 4 \cos \left(\frac{1}{4}\pi \right) = 4 \left(\frac{\sqrt{2}}{2} \right) = 2\sqrt{2};$$

es decir, las coordenadas cartesianas del punto con coordenadas esféricas $(4, \frac{2}{3}\pi, \frac{1}{4}\pi)$, serán $(-\sqrt{2}, \sqrt{6}, 2\sqrt{2})$.

Las coordenadas esféricas se usan en problemas de simetría esférica, simetría alrededor de un punto, mientras que las coordenadas cilíndricas se usan en problemas de simetría cilíndricas, simetría alrededor de una recta.

Así como hay vectores unitarios asociados a las coordenadas cartesianas, también hay vectores unitarios asociados a las coordenadas esféricas y cilíndricas, que veremos más adelante.

Ejercicios.

- 1) Demuestra que la distancia entre dos puntos $P_1 (r_1, \theta_1)$ y $P_2 (r_2, \theta_2)$ en coordenadas polares es

$$d(P_1, P_2) = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos(\theta_1 - \theta_2)}.$$

- 2) Representa las curvas que tienen las siguientes ecuaciones en coordenadas polares.

a) $r = \frac{3}{\text{sen } \theta - \cos \theta};$

b) $r = \frac{4}{4 - \cos \theta};$

c) $r = \frac{3}{2 - \text{sen } \theta};$

d) $r = \frac{1}{1 + 2 \cos \theta};$

e) $r = \frac{1}{\theta}, \quad 0 < \theta \leq 2\pi.$

- 3) Halla las coordenadas cilíndricas y esféricas de las siguientes superficies que están en coordenadas rectangulares cartesianas.

a) $x^2 + y^2 - 4z^2 = 16;$

b) $x^2 + y^2 = 9z;$

c) $x^2 + y^2 + z^2 - 2y = 0.$

- 4) Representa las superficies cuyas ecuaciones en coordenadas cilíndricas son

a) $r^2 + z^2 = 9;$

b) $r = z;$

c) $r = 2 \cos \theta.$

- 5) Describe en coordenadas cilíndricas

9.5. ESPACIO EUCLÍDEO DE DIMENSIÓN N .

- a) la región del espacio de los puntos que están a una distancia a del Polo y se encuentran en el interior del cilindro de ecuación $r = a \cos \theta$;
- b) el cono sólido de altura y radio iguales a 4 unidades.
- 6) Representa las siguientes regiones descritas por las siguientes restricciones en coordenadas cilíndricas.
- a) $0 \leq \theta \leq \pi/2 \quad 1 \leq z \leq 2 \quad 1 \leq r \leq z$;
- b) $0 \leq \theta \leq \pi/2 \quad 0 \leq r \leq \cos \theta \quad 1 \leq z \leq 2$.
- 7) a) ¿ Cuáles son las coordenadas cilíndricas del punto que en coordenadas esféricas tiene por coordenadas (ρ, θ, ϕ) ?
- b) Representa el conjunto de puntos (ρ, θ, ϕ) tales que $\phi = \theta = \pi/2$.
- 8) Describe las siguientes regiones del espacio \mathbb{R}^3 utilizando coordenadas esféricas.
- a) la intersección del cono con vértice en el Polo, eje de simetría el eje z y ángulo entre el eje de simetría y una generatriz igual a $\pi/6$ radianes, con la esfera de radio a y centro el Polo;
- b) la región comprendida entre las esferas centradas en el origen y con radios a y b , respectivamente $a < b$.
- 9) Representa las regiones descritas por las siguientes restricciones en coordenadas esféricas.
- a) $0 \leq \theta \leq \pi/2 \quad 0 \leq \phi \leq \pi/2 \quad 0 \leq \rho \leq 1$;
- b) $0 \leq \theta \leq \pi/2 \quad 0 \leq \phi \leq \pi/4 \quad 1 \leq \rho \leq 2 \sec \phi$.
- 10) Interpreta geoméricamente cada una de las ecuaciones siguientes dadas en coordenadas esféricas.
- a) $\rho \sin \phi = 1$;
- b) $\sin \phi = 1$;
- c) $\cos \phi = -\frac{1}{2} \sqrt{2}$;
- d) $\tan \theta = 1$;
- e) $\rho \cos \phi = 1$;
- f) $\rho = \cos \phi$.

9.5 Espacio euclídeo de dimensión n .

En esta última sección vamos a extender los conceptos de vectores, ya estudiados en \mathbb{R}^2 y \mathbb{R}^3 , a un nuevo espacio, \mathbb{R}^n , llamado n -espacio euclidiano. Hemos visto que podemos pensar \mathbb{R}^3 de cualquiera de estas dos maneras:

- i) algebraicamente, como un conjunto de ternas (x, y, z) de números reales;
- ii) geoméricamente, como un conjunto de segmentos dirigidos, que hemos llamados vectores.

La primera manera es más conveniente para generalizar.

Definición. Se define el **espacio euclidiano n -dimensional** \mathbb{R}^n como el conjunto de todas las n -adas ordenadas (x_1, x_2, \dots, x_n) , donde los x_i son números reales. Los elementos $\vec{x} = (x_1, x_2, \dots, x_n)$ se conocen como **n -vectores**.

Obviamente, al poner $n = 1, 2$ o 3 recuperamos la recta, el plano y el espacio tridimensional, respectivamente.

Por extensión a lo ya dicho, tenemos en este espacio las siguientes dos operaciones:

1. **Suma:**

$$\vec{x} + \vec{y} = (x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n);$$

2. **Multipliación por un escalar:**

$$\alpha \vec{x} = \alpha (x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

Los n vectores $\vec{e}_1 = (1, 0, \dots, 0)$, $\vec{e}_2 = (0, 1, 0, \dots, 0)$, \dots , $\vec{e}_n = (0, 0, \dots, 0, 1)$, son los **vectores de la base usual de \mathbb{R}^n** , y generalizan los tres vectores unitarios ortogonales entre sí, $\vec{i}, \vec{j}, \vec{k}$, de la base de \mathbb{R}^3 .

Un vector genérico de \mathbb{R}^n , $\vec{x} = (x_1, x_2, \dots, x_n)$, se puede escribir como

$$\vec{x} = (x_1, x_2, \dots, x_n) = x_1 \vec{e}_1 + x_2 \vec{e}_2 + \dots + x_n \vec{e}_n.$$

Dados dos vectores $\vec{x} = (x_1, x_2, \dots, x_n)$ y $\vec{y} = (y_1, y_2, \dots, y_n)$ se extiende el **producto escalar** de la siguiente manera:

$$\vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Continuando con la analogía con \mathbb{R}^3 , definimos la **norma** de un vector $\vec{x} \in \mathbb{R}^n$ como

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Si \vec{x} e \vec{y} son dos vectores en el plano o en el espacio, sabemos que el ángulo θ entre ellos está dado por la fórmula dada en el teorema 9.1

$$\cos \theta = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}.$$

Esto sigue siendo válido también para vectores de \mathbb{R}^n , dado que \vec{x} e \vec{y} están en un subespacio bidimensional de \mathbb{R}^n (es decir, un plano) y el ángulo entre ellos está bien definido.

Como en el caso de vectores en el plano o en el espacio, son válidas las siguientes propiedades, para $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ y $\alpha, \beta \in \mathbb{R}$.

1. $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$;

9.5. ESPACIO EUCLÍDEO DE DIMENSIÓN N .

2. $(\alpha \vec{x} + \beta \vec{y}) \cdot \vec{z} = \alpha (\vec{x} \cdot \vec{z}) + \beta (\vec{y} \cdot \vec{z})$;
3. $\vec{x} \cdot \vec{x} \geq 0$; y $\vec{x} \cdot \vec{x} = 0$ si y sólo si $\vec{x} = \vec{0}$.

En el caso de vectores de \mathbb{R}^3 , probamos una propiedad mucho más interesante del producto escalar, llamada la desigualdad de Cauchy-Schwartz (ver corolario 9.2), usando la ley de los cosenos. En el caso de vectores en \mathbb{R}^n podemos dar una demostración algebraica.

Teorema 9.3 (Desigualdad de Cauchy-Schwartz).

Sean \vec{x} e \vec{y} vectores en \mathbb{R}^n . Entonces

$$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|.$$

Demostración:

Sea $a = \vec{y} \cdot \vec{y}$ y $b = -\vec{x} \cdot \vec{y}$. Si $a = 0$, entonces $\vec{y} = \vec{0}$ y ambos lados se reducen a cero, dando la validez del teorema. Podemos entonces suponer que sea $a \neq 0$; por las propiedades dadas arriba, tenemos

$$\begin{aligned} 0 &\leq (a\vec{x} + b\vec{y}) \cdot (a\vec{x} + b\vec{y}) = a^2 (\vec{x} \cdot \vec{x}) + 2ab (\vec{x} \cdot \vec{y}) + b^2 (\vec{y} \cdot \vec{y}) = \\ &= (\vec{y} \cdot \vec{y})^2 (\vec{x} \cdot \vec{x}) - 2 (\vec{y} \cdot \vec{y}) (\vec{x} \cdot \vec{y}) (\vec{x} \cdot \vec{y}) + (\vec{x} \cdot \vec{y})^2 (\vec{y} \cdot \vec{y}) = \\ &= (\vec{y} \cdot \vec{y})^2 (\vec{x} \cdot \vec{x}) - (\vec{x} \cdot \vec{y})^2 (\vec{y} \cdot \vec{y}), \end{aligned}$$

dividiendo por $\vec{y} \cdot \vec{y}$ (que es siempre positivo y supuestamente no nulo), se tiene

$$0 \leq (\vec{y} \cdot \vec{y}) (\vec{x} \cdot \vec{x}) - (\vec{x} \cdot \vec{y})^2 \iff (\vec{x} \cdot \vec{y})^2 \leq (\vec{y} \cdot \vec{y}) (\vec{x} \cdot \vec{x}) = \|\vec{x}\|^2 \|\vec{y}\|^2.$$

Al extraer la raíz cuadrada se obtiene la regla deseada.

c.q.d.

Como consecuencia de la desigualdad de Cauchy-Schwartz, tenemos el siguiente corolario.

Corolario 9.4 (Desigualdad triangular).

Sean \vec{x} e \vec{y} vectores en \mathbb{R}^n . Entonces

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|.$$

Demostración:

Por el teorema 9.3, $\vec{x} \cdot \vec{y} \leq |\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$ de modo que

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= (\vec{x} + \vec{y}) \cdot (\vec{x} + \vec{y}) = \vec{x} \cdot \vec{x} + 2\vec{x} \cdot \vec{y} + \vec{y} \cdot \vec{y} = \|\vec{x}\|^2 + 2\vec{x} \cdot \vec{y} + \|\vec{y}\|^2 \leq \\ &\leq \|\vec{x}\|^2 + 2\|\vec{x}\| \|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2; \end{aligned}$$

al extraer la raíz cuadrada se obtiene el resultado.

c.q.d.

Por analogía con \mathbb{R}^3 , podemos definir la **distancia** en \mathbb{R}^n . $\|\vec{x} - \vec{y}\|$ es la longitud del vector $\vec{x} - \vec{y}$ y nos da la distancia entre \vec{x} e \vec{y} .

Lo que **no se puede generalizar** es el producto vectorial; este sólo está definido para vectores en el espacio tridimensional.

Ejercicios.

- 1) Prueba las propiedades del producto escalar para vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$, esto es que $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$, $\vec{x} \cdot \vec{x} \geq 0$ y $\vec{x} \cdot \vec{x} = 0$ si y sólo si $\vec{x} = \vec{0}$.
- 2) Verifica que se cumplen las desigualdades de Cauchy-Schwartz y triangular, para los vectores
 - a) $\vec{x} = (2, 0, -1)$ y $\vec{y} = (4, 0, -2)$;
 - b) $\vec{x} = (1, 0, 2, 6)$ y $\vec{y} = (3, 8, 4, 1)$;
 - c) $\vec{x} = (1, -1, 1, -1, 1)$ y $\vec{y} = (3, 0, 0, 0, 2)$.

Capítulo 10

Diferenciación.

10.1 La geometría de las funciones con valores reales.

Definición. Sea \vec{f} una función cuyo dominio sea un subconjunto A de \mathbb{R}^n y cuya imagen esté contenida en \mathbb{R}^m :

$$\vec{f}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m;$$

es decir, a cada elemento $\vec{x} = (x_1, x_2, \dots, x_n) \in A$, \vec{f} asigna un valor $\vec{f}(\vec{x})$ que será un elemento de \mathbb{R}^m . Dichas funciones \vec{f} se llaman **funciones con valores vectoriales** si $m > 1$, y **funciones con valores reales**, o escalares, si $m = 1$. Si $n > 1$, se llaman **funciones de varias variables** en contraposición a las funciones de una variable real ($n = 1$): $\vec{f}: A \subset \mathbb{R} \rightarrow \mathbb{R}^m$. Cuando $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ se dice que f es una **función de n variables, con dominio U y valores reales**.

Las funciones de varias variables surgen de manera natural en contextos muy elementales, como por ejemplo:

1. $f(x, y) = \sqrt{x^2 + y^2}$ da la distancia desde un punto $P(x, y)$ hasta el origen en el plano;
2. $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ da la distancia desde un punto $P(x, y, z)$ hasta el origen en el espacio tridimensional;
3. $f(x, y, z) = xyz$ da el volumen de un paralelepípedo rectangular de dimensiones x, y, z ;
4. Para especificar la temperatura T en una región del espacio se requiere una función $T: A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, de manera que $T(x, y, z)$ sea la temperatura en el punto (x, y, z) ;
5. Para especificar la velocidad de un fluido moviéndose en el espacio se requiere una asociación $\vec{V}: A \subset \mathbb{R}^4 \rightarrow \mathbb{R}^3$, de manera que $\vec{V}(x, y, z, t)$ sea el vector velocidad del fluido en el punto (x, y, z) del espacio y en el tiempo t .

Si el dominio de una función de varias variables no viene dado explícitamente, se sobreentiende que el dominio es el conjunto de todos los puntos para los cuales la definición tiene sentido. Como en el caso de funciones reales de una variable real, decir que la función está **acotada** es decir que su imagen está acotada. Así, por ejemplo, en el caso de $f(x, y) = \frac{1}{x - y}$ se entiende que el dominio está formado por todos los puntos (x, y) tales que $x \neq y$. Dado que la función toma todos los valores distintos de cero, esta función no está acotada.

Como ya sabemos, para $f: U \subset \mathbb{R} \rightarrow \mathbb{R}$, la **gráfica** de f es el subconjunto de \mathbb{R}^2 formado por los puntos $(x, f(x))$ en el plano, para $x \in U$. Este subconjunto se puede interpretar como una curva de \mathbb{R}^2 . Esta idea se puede generalizar para funciones de varias variables.

Definición. Sea $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función de n variables con valores reales. Definimos la **gráfica** de f como el subconjunto de \mathbb{R}^{n+1} formado por los puntos $(\vec{x}, f(\vec{x})) = (x_1, x_2, \dots, x_n, f(x_1, x_2, \dots, x_n))$ de \mathbb{R}^{n+1} , para $\vec{x} = (x_1, x_2, \dots, x_n) \in U$. En símbolos

$$\text{gráfica de } f = \{(x_1, x_2, \dots, x_n, f(x_1, x_2, \dots, x_n)) \in \mathbb{R}^{n+1} \mid (x_1, x_2, \dots, x_n) \in U\}.$$

Para el caso $n = 1$, la gráfica es una curva en \mathbb{R}^2 , mientras que para $n = 2$ es una superficie en \mathbb{R}^3 . Para $n = 3$ sería un subconjunto de \mathbb{R}^4 difícil de imaginar. Para superar esta dificultad, introducimos la idea de conjunto de nivel.

Definición. Sea $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y sea $c \in \mathbb{R}$. Entonces el **conjunto de nivel del valor** c se define como aquellos puntos $\vec{x} \in U$ para los cuales $f(\vec{x}) = c$. Si $n = 2$, hablamos de una **curva de nivel** (de valor c); si $n = 3$, hablamos de una **superficie de nivel**. En símbolos, el conjunto de nivel de valor c se escribe

$$\{(x_1, x_2, \dots, x_n) \in U \mid f(x_1, x_2, \dots, x_n) = c\} \subset \mathbb{R}^n.$$

Nótese que el conjunto de nivel siempre está en el espacio dominio.

Es posible determinar el aspecto de una gráfica mediante el método de las secciones; con más precisión, hablaremos de **trazas** cuando hagamos la intersección de la gráfica con los planos coordenados y de **secciones** cuando se hace la intersección de la gráfica con los planos en general. Nótese que las secciones con planos paralelos al plano xy no son nada más que las curvas de nivel.

Visto que más adelante nos encontraremos, una y otra vez, con algunas superficies especiales, vamos a estudiar con detalle las superficies asociadas a ecuaciones de la forma

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Hx + Iy + Jz + K = 0.$$

Tales superficies se denominan **superficies cuádricas**. Mediante traslaciones y rotaciones convenientemente elegidas de los ejes coordenados, podemos simplificar tales ecuaciones y demostrar que las cuádricas no degeneradas pertenecen a uno de los nueve tipos siguientes: elipsoide, hiperboloide de una hoja, hiperboloide de dos hojas, cono, paraboloides elíptico, paraboloides hiperbólico, cilindro parabólico, cilindro elíptico, cilindro

10.1. LA GEOMETRÍA DE LAS FUNCIONES CON VALORES REALES.

hiperbólico. Nótese que excluimos la cuádricas degeneradas como $1 + x^2 + y^2 = -z^2$ y $x^2 + y^2 = -z^2$. Aquí daremos una descripción de cada una de ellas, junto con su ecuación en forma estándar y algunas de sus propiedades especiales. Las primeras a considerar son las superficies en cuyas ecuaciones en forma estándar aparecen x^2, y^2, z^2 .

1. El elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

El elipsoide está centrado en el origen y es simétrico con respecto a los tres planos coordenados.

Corta los ejes coordenados en seis puntos $(\pm a, 0, 0)$, $(0, \pm b, 0)$, $(0, 0, \pm c)$, que se llaman los vértices.

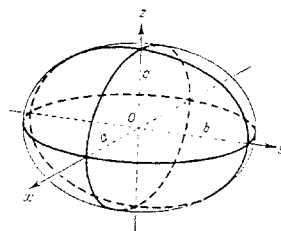
La superficie está acotada al estar contenida en el paralelepípedo: $|x| \leq a$, $|y| \leq b$, $|z| \leq c$.

Las tres trazas son elipses: así, por ejemplo, la traza sobre el plano xy (hacer $z = 0$) es la elipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Todas las secciones paralelas a los planos coordenados también son elipses, por ejemplo, tomando $y = y_0$ tenemos $\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1 - \frac{y_0^2}{b^2}$.

Los números a, b, c se llaman los semiejes del elipsoide. Si dos de los semiejes son iguales tenemos un elipsoide de revolución: si los tres semiejes son iguales, la superficie es una esfera.



2. El hiperboloide de una hoja

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

La superficie no está acotada.

Está centrada en el origen y es simétrica con respecto a los tres planos coordenados.

La superficie corta los ejes coordenados en cuatro puntos $(\pm a, 0, 0)$, $(0, \pm b, 0)$.

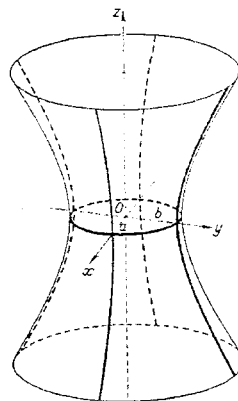
La traza sobre el plano xy (poner $z = 0$) es la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

Todas las secciones paralelas al plano xy son elipses.

La traza en el plano xz (hacer $y = 0$) es la hipérbola $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$ y la traza en el plano yz (hacer $x = 0$) es la hipérbola $\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.

La traza en el plano xz (hacer $y = 0$) es la hipérbola $\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1$ y la traza en el plano yz (hacer $x = 0$) es la hipérbola $\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.

es la hipérbola $\frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.



Todas las secciones paralelas a los planos xz e yz son hipérbolas. Si $a = b$, todas las secciones paralelas al plano xy son circunferencias y tenemos un hiperboloide de revolución.

3. El hiperboloide de dos hojas

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1.$$

La superficie corta los ejes coordenados sólo en los dos vértices $(0, 0, \pm c)$.

La superficie consta de dos partes: una para la cual $z \geq c$ y otra para la cual $z \leq -c$.

Todas las secciones paralelas al plano xy son elipses: para $z = z_0$ con $|z_0| \geq c$, tenemos

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z_0^2}{c^2} - 1.$$

Las secciones paralelas a los demás planos coordenados son hipérbolas; por ejemplo, para $y = y_0$, tenemos

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} = 1 + \frac{y_0^2}{b^2}.$$

La superficie es simétrica con respecto a los tres planos coordenados y está centrada en el origen.

4. El cono

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z^2.$$

La superficie sólo corta los ejes coordenados en el origen.

La superficie no está acotada.

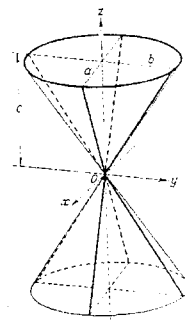
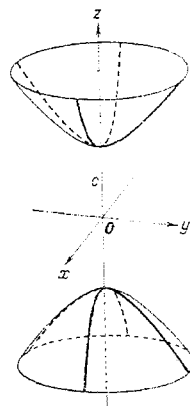
Hay simetría con respecto a los tres planos coordenados.

La traza sobre el plano xz ($y = 0$) es un par de rectas que se cortan $z = \pm \frac{x}{a}$.

La traza en el plano yz ($x = 0$) también es un par de rectas que se cortan $z = \pm \frac{y}{b}$.

La traza en el plano xy ($z = 0$) se reduce al origen.

Las secciones paralelas al plano xy son elipses. Si $a = b$, estas secciones son circunferencias y tenemos lo que comúnmente se llama cono circular doble (o simplemente cono). Las dos porciones superior e inferior del cono se llaman hojas.



10.1. LA GEOMETRÍA DE LAS FUNCIONES CON VALORES REALES.

Pasamos ahora a los paraboloides. En las ecuaciones en forma estándar aparecen x^2 , y^2 , pero en lugar de z^2 aparece z .

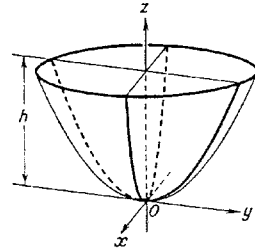
5. El paraboloide elíptico

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z.$$

La superficie no se extiende por debajo del plano xy .

Además no está acotada. El origen se denomina vértice.

Las secciones paralelas al plano xy son elipses; las secciones paralelas a los demás planos coordenados son parábolas. De ahí el término paraboloide elíptico.



La superficie es simétrica con respecto al plano xz y al plano yz . También es simétrica con respecto al eje z . Si $a = b$, la superficie es un paraboloide de revolución.

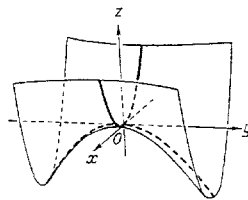
6. El paraboloide hiperbólico

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z.$$

Aquí también hay simetría con respecto al plano xz y al plano yz .

Las secciones paralelas al plano xy son hipérbolas; las secciones paralelas a los demás planos coordenados son parábolas. De ahí el término paraboloide hiperbólico. Si lo miramos al revés es como "una silla de montar a caballo".

El origen es un punto de mínimo para la traza en el plano xz , pero es un punto de máximo para la traza en el plano yz .



El origen se llama minimax o punto de ensilladura de la superficie.

El resto de las superficies cuádricas son cilindros, cuya definición es la siguiente: tómesese cualquier curva plana C , todas las rectas que pasan por C y son perpendiculares al plano de C forman una superficie que se llama cilindro de base la curva C . Las rectas perpendiculares se conocen como generatrices del cilindro. Si la curva de base está en el plano xy , o en un plano paralelo, las generatrices del cilindro son paralelas al eje z : en la ecuación de la superficie sólo aparecen x e y . No hay restricciones en cuanto a la coordenada z , que puede tomar todos los valores. Claramente estas superficies no están acotadas.

7. El cilindro parabólico

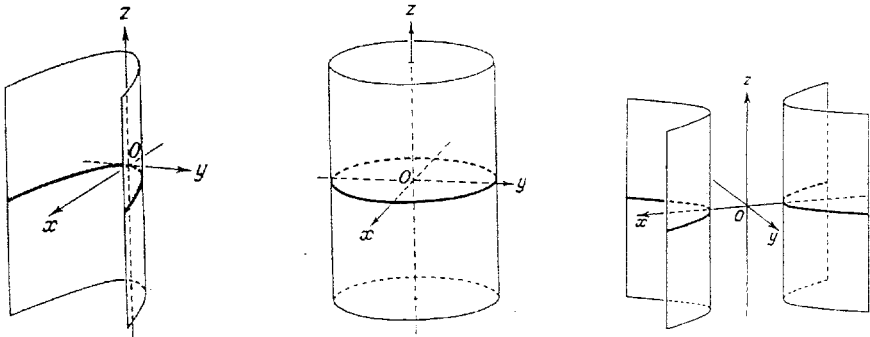
$$x^2 = ay.$$

Esta superficie está formada por todas las rectas que pasan por la parábola $x^2 = ay$ y son perpendiculares al plano xy .

8. El cilindro elíptico

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Esta superficie está formada por todas las rectas que pasan por la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ y son perpendiculares al plano xy . Si $a = b$, tenemos el cilindro circular recto habitual.



9. El cilindro hiperbólico

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

La superficie consta de dos partes, cada una de ellas generadas por una rama de la hipérbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

Veamos ahora como construir las gráficas de algunas otras funciones interesantes.

Ejemplo 1. Describe la gráfica de la función

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

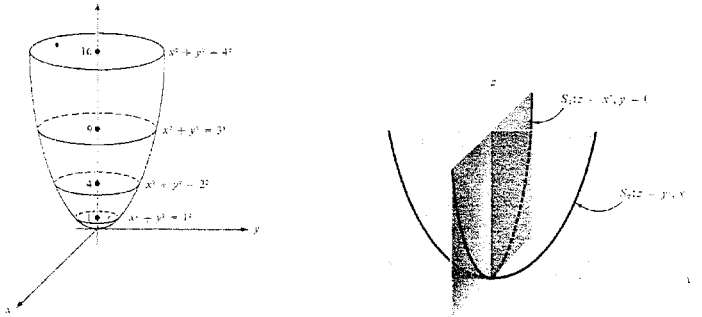
Solución:

El dominio es todo el plano \mathbb{R}^2 . La gráfica de f es un paraboloides de revolución del tipo dado arriba. Esta superficie puede ser generada al hacer girar la parábola $z = x^2$ alrededor del eje z .

La curva de nivel del valor c no existe si $c < 0$; para $c > 0$ es el conjunto de los puntos del plano $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c\}$, es decir un círculo de radio \sqrt{c} con centro en el origen, que para $c = 0$ se reduce solamente al punto origen.

10.1. LA GEOMETRÍA DE LAS FUNCIONES CON VALORES REALES.

La traza en el plano xz , es decir poniendo $y = 0$, será la parábola $z = x^2$, y la traza en el plano yz , es decir poniendo $x = 0$, será la parábola $z = y^2$.



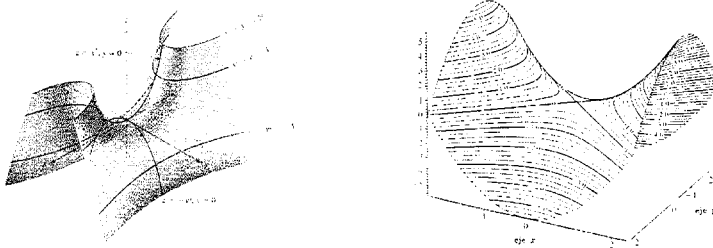
Ejemplo 2. Describe la gráfica de la función

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

Solución:

El dominio es todo el plano \mathbb{R}^2 . La gráfica de f es un paraboloides hiperbólico del tipo dado arriba.

Estudiemos las curvas de nivel resolviendo la ecuación $x^2 - y^2 = c$, considerando los valores $c = 0, \pm 1, \pm 4$. Para $c = 0$ tenemos $y^2 = x^2$, es decir la curva de nivel está formada por dos rectas que pasan por el origen, $y = \pm x$. Para $c = 1$, la curva de nivel es $x^2 - y^2 = 1$, o sea $y = \pm\sqrt{x^2 - 1}$ que es una hipérbola que cruza verticalmente el eje x en los puntos $(\pm 1, 0)$. De manera análoga, para $c = 4$, la curva de nivel es $x^2 - y^2 = 4$, o sea $y = \pm\sqrt{x^2 - 4}$ que es una hipérbola que cruza verticalmente el eje x en los puntos $(\pm 2, 0)$. Para $c = -1$, la curva de nivel es $x^2 - y^2 = -1$, o sea $x = \pm\sqrt{y^2 - 1}$ que es una hipérbola que cruza horizontalmente el eje y en los puntos $(0, \pm 1)$. Y para $c = -4$, la curva de nivel es $x^2 - y^2 = -4$, o sea $x = \pm\sqrt{y^2 - 4}$ que es una hipérbola que cruza horizontalmente el eje y en los puntos $(0, \pm 2)$.



Como no es fácil visualizar la gráfica de f a partir sólo de estos datos, calculamos las otras dos trazas. La traza en el plano xz , es decir poniendo $y = 0$, será la parábola que se abre hacia arriba $z = x^2$; y la traza en el plano yz , es decir poniendo $x = 0$, será la parábola que se abre hacia abajo $z = -y^2$.

Ejemplo 3. Describe la gráfica de la función

$$f(x, y) = \sqrt{r^2 - (x^2 + y^2)}.$$

Solución:

Esta función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ está definida en el disco cerrado $x^2 + y^2 \leq r^2$. Estudiemos las curvas de nivel resolviendo la ecuación $\sqrt{r^2 - (x^2 + y^2)} = c$. La curva de nivel del valor c no existe si $c < 0$; para $c = 0$ la curva de nivel se reduce al círculo de radio r y centro el origen, mientras que para $0 < c < r$ es el conjunto $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2 - c^2\}$, es decir un círculo de radio $\sqrt{r^2 - c^2}$ con centro en el origen. Si $c = r$ la curva de nivel se reduce al único punto $(0, 0)$. No hay curva de nivel si $c > r$. La gráfica es la mitad superior de la esfera.

Ejemplo 4. Describe la gráfica de la función

$$f(x, y) = x + y + 1.$$

Solución:

La función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ está definida en todo \mathbb{R}^2 . Tiene como gráfica el plano inclinado $z = x + y + 1$. Este plano intersecta el plano xy ($z = 0$) en la recta $y = -x - 1$ y el eje z en el punto $(0, 0, 1)$. Para cualquier valor de $c \in \mathbb{R}$, la curva de nivel del valor c es la recta $y = -x + (c - 1)$.

Ejemplo 5. Describe la gráfica de la función

$$f(x, y, z) = Ax + By + Cz.$$

Solución:

La función $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ está definida en todo \mathbb{R}^3 . La gráfica está en el espacio \mathbb{R}^4 que no sabemos imaginar; sin embargo sí podemos visualizar las superficies de nivel, que resultan ser los planos paralelos $Ax + By + Cz = c$.

Ejemplo 6. Describe la gráfica de la función

$$f(x, y, z) = \sqrt{x^2 + y^2 + z^2}.$$

Solución:

La función $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ está definida en todo \mathbb{R}^3 . La gráfica está en el espacio \mathbb{R}^4 que no sabemos imaginar; sin embargo sí podemos visualizar las superficies de nivel, que resultan ser las esferas concéntricas $x^2 + y^2 + z^2 = c^2$; tienen centro en el origen $(0, 0, 0)$ y radio c , para $c > 0$. No hay superficies de nivel para $c < 0$. Como ya sabemos se obtiene mayor información acerca de la gráfica de una función calculando una sección. Por ejemplo, si consideramos el equivalente del plano xy ($z = 0$) en el caso cuatridimensional, podemos ver la sección de la gráfica de f : $\{(x, y, z, t) \in \mathbb{R}^4 \mid t = x^2 + y^2, z = 0\}$, que se encuentra en \mathbb{R}^3 , y que resulta ser un paraboloides de revolución.

Ejemplo 7. Describe la gráfica de la función

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

10.1. LA GEOMETRÍA DE LAS FUNCIONES CON VALORES REALES.

Solución:

La función $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ está definida en todo \mathbb{R}^3 . La gráfica está en el espacio \mathbb{R}^4 que no sabemos imaginar; sin embargo sí podemos visualizar las superficies de nivel, que resultan ser las superficies dadas por $x^2 + y^2 - z^2 = c$. Para $c = 0$, se trata de un cono $z = \pm\sqrt{x^2 + y^2}$. Para $c > 0$, digamos $c = a^2$, la superficie de nivel es el hiperboloide de revolución de una hoja alrededor del eje z , definido por $z = \pm\sqrt{x^2 + y^2 - a^2}$, el cual intersecta el plano xy en un círculo de radio $|a|$ y centro $(0, 0)$. Para $c < 0$, digamos $c = -b^2$, la superficie de nivel es el hiperboloide de dos hojas alrededor del eje z , definido por $z = \pm\sqrt{x^2 + y^2 + b^2}$, que atraviesa el eje z en los puntos $(0, 0, \pm b)$. Se obtiene mayor información acerca de la gráfica de una función calculando una sección. Por ejemplo, si consideramos el equivalente del plano xz ($y = 0$) en el caso cuadrimensional, podemos ver la sección de la gráfica de $f: \{(x, y, z, t) \in \mathbb{R}^4 / t = x^2 - z^2, y = 0\}$, que se encuentra en \mathbb{R}^3 , y que resulta ser un paraboloides hiperbólico, es decir la silla de montar.

Ejercicios.

- 1) Halla el valor de la función $f(x, y) = \frac{x^4 + 2x^2y^2 + y^4}{1 - x^2 - y^2}$ en los puntos de la circunferencia $x^2 + y^2 = R^2$.
- 2) Halla y representa los campos de existencia de las siguientes funciones:
 - a) $f(x, y) = 1 + \sqrt{1 - (x - y)^2}$;
 - b) $f(x, y) = \sqrt{x^2 - 4} + \sqrt{4 - y^2}$;
 - c) $f(x, y) = \sqrt{x^2 + y^2 - a^2} \cdot \sqrt{2a^2 - x^2 - y^2}$ con $a > 0$;
 - d) $f(x, y) = (y \operatorname{sen} x)^{1/2}$;
 - e) $f(x, y) = \log(x^2 + y)$.
- 3) Halla los campos de existencia de las funciones de tres variables:
 - a) $f(x, y, z) = \log(xyz)$;
 - b) $f(x, y, z) = \sqrt{1 - x^2 - y^2 - z^2}$;
 - c) $f(x, y, z) = \arcsen x + \arcsen y + \arcsen z$.
- 4) Identifica las curvas de nivel y dibuja algunas de ellas.
 - a) $f(x, y) = \exp(xy)$;
 - b) $f(x, y) = x^3 - y$;
 - c) $f(x, y) = x^2 - y^2$;
 - d) $f(x, y) = x^2y^2$;
 - e) $f(x, y) = \frac{x^2}{x^2 + y^2}$.
- 5) Identifica las superficies de nivel para las siguientes funciones.
 - a) $f(x, y, z) = x + y + z$ para $c = -1, 0, 1$;
 - b) $f(x, y, z) = x^2 + 2y^2 + 3z^2$ para $c = 0, 6, 12$;
 - c) $f(x, y, z) = x + 2y + 3z$ para $c = 0$.

10.2 Límites y continuidad.

En esta sección veremos de forma superficial en los conceptos de conjunto abierto, entorno, límite y continuidad: los conjuntos abiertos y entornos son necesarios para entender límites, y, a su vez, los límites son necesarios para entender continuidad y diferenciabilidad. Sin embargo, como sucede en el cálculo elemental, no es necesario dominar completamente el concepto de límite para poder trabajar con problemas de diferenciación, y por lo tanto no seremos del todo rigurosos. Empezamos con la definición de entorno y de conjunto abierto.

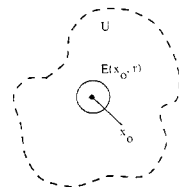
Definición. Un **entorno** de un punto $\vec{x}_0 \in \mathbb{R}^n$ es un conjunto de la forma $E(\vec{x}_0, r) = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x} - \vec{x}_0\| < r\}$ donde r es un número real positivo. Este conjunto se llama también disco abierto (o bola abierta) de radio r y centro \vec{x}_0 , y es el conjunto de todos los puntos de \mathbb{R}^n cuya distancia a \vec{x}_0 es menor que r . Diremos que U es un **entorno perforado** de \vec{x}_0 si es un conjunto de la forma $\{\vec{x} \in \mathbb{R}^n \mid 0 < \|\vec{x} - \vec{x}_0\| < r\}$; es decir, se le ha quitado el centro.

En el caso $n = 1$, el entorno de $x_0 \in \mathbb{R}$ es el intervalo abierto $(x_0 - r, x_0 + r)$. En el caso $n = 2$, el entorno de $\vec{x}_0 \in \mathbb{R}^2$ es el *interior* del disco de radio r con centro en \vec{x}_0 . En el caso $n = 3$, el entorno de $\vec{x}_0 \in \mathbb{R}^3$ es el *interior* de la esfera de radio r con centro en \vec{x}_0 .

Definición. Sea $U \subset \mathbb{R}^n$ un subconjunto de \mathbb{R}^n . Diremos que un punto \vec{x}_0 es un **punto interior** del conjunto U , si el conjunto U contiene algún entorno $E(\vec{x}_0, r)$ de \vec{x}_0 . El conjunto de todos los puntos interiores se llama **interior** de U .

Definición. Diremos que un punto \vec{x}_0 es un **punto de frontera** del conjunto U , si y sólo si todo entorno $E(\vec{x}_0, r)$ de \vec{x}_0 contiene puntos que están en U y puntos que no están en U . El conjunto de todos los puntos de frontera se llama **frontera** de U , y se denota como ∂U .

Definición. U es un **conjunto abierto** cuando para cualquier punto $\vec{x}_0 \in U$ existe un entorno $E(\vec{x}_0, r)$ de \vec{x}_0 contenido en U ; en símbolos $E(\vec{x}_0, r) \subset U$. Es decir, un conjunto es abierto si y sólo si cada uno de sus puntos es punto interior. Nótese que la magnitud $r > 0$ depende del punto \vec{x}_0 , en general r disminuirá conforme \vec{x}_0 se acerca al borde de U .



Además, por convenio, el conjunto vacío \emptyset es abierto.

Definición. Diremos que el conjunto U es **cerrado** si y sólo si contiene su frontera ∂U .

Ejemplo 1. Sean los conjuntos

$$S_1 = \{(x, y) \in \mathbb{R}^2 \mid 1 < x < 2, 1 < y < 2\},$$

$$S_2 = \{(x, y) \in \mathbb{R}^2 \mid 3 \leq x \leq 4, 1 \leq y \leq 2\}.$$

$$S_3 = \{(x, y) \in \mathbb{R}^2 \mid 5 \leq x \leq 6, 1 < y < 2\}.$$

10.2. LÍMITES Y CONTINUIDAD.

Estos conjuntos son superficies cuadradas en el plano. El primero, S_1 , es abierto, ya que es sólo el interior del cuadrado y contiene un entorno de cada uno de sus puntos. S_2 es cerrado, ya que está formado por el interior del cuadrado junto con los cuatro lados, es decir contiene toda su frontera. Finalmente, S_3 está formado por el interior del cuadrado junto con sus dos lados verticales; luego S_3 no es abierto ya que contiene parte de su frontera, y no es cerrado por que no contiene toda su frontera.

Ejemplo 2. Sean los conjuntos

$$S_1 = \{(x, y, z) \in \mathbb{R}^3 \mid z > x^2 + y^2\},$$

$$S_2 = \{(x, y, z) \in \mathbb{R}^3 \mid z \geq x^2 + y^2\},$$

$$S_3 = \{(x, y, z) \in \mathbb{R}^3 \mid 1 \geq \frac{x^2 + y^2}{z}\}.$$

La frontera de estos tres conjuntos es el paraboloides de revolución $z = x^2 + y^2$.

El primer conjunto consta de todos los puntos por encima de esta superficie y es un conjunto abierto. El segundo conjunto es cerrado ya que contiene toda su frontera. El tercer conjunto, ni es abierto, ni es cerrado. No es abierto porque contiene algunos de sus puntos frontera; por ejemplo, contiene el punto $(1, 1, 2)$. No es cerrado porque no contiene el punto frontera $(0, 0, 0)$.

Una vez introducidos estos conceptos, podemos pasar a definir límites. Vamos a suponer de aquí en adelante que el dominio de definición de la función f es un conjunto abierto A . Estaremos interesados en hallar el límite de f cuando $x \in A$ tienda a un punto de A o de su frontera. El concepto de límite sirve para: derivadas, máximos, mínimos, asíntotas, integrales impropias, series y sucesiones, etc.

Vamos a dar el concepto de límite para funciones de varias variables. Por la teoría de funciones de una variable, la definición de $\lim_{x \rightarrow x_0} f(x) = l$ para $f: A \subset \mathbb{R} \rightarrow \mathbb{R}$ es ya conocida. Intuitivamente, esto significa que conforme x se acerca a x_0 , $f(x)$ se acerca a l . Lo mismo sucede con funciones de varias variables. El enfoque más intuitivo es el enfoque con entornos.

Definición. Sea $\vec{f}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, donde A es un conjunto abierto. Sea \vec{x}_0 un punto en A o en la frontera de A , y sea $E(\vec{l}, r)$ un entorno de $\vec{l} \in \mathbb{R}^m$. Decimos que $\vec{f}(\vec{x})$ **tiende a** \vec{l} cuando \vec{x} tiende a \vec{x}_0 , o, en símbolos,

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{l} \quad \text{o} \quad \vec{f}(\vec{x}) \rightarrow \vec{l} \quad \text{cuando} \quad \vec{x} \rightarrow \vec{x}_0$$

cuando, dado cualquier entorno $E(\vec{l}, \epsilon)$ de \vec{l} , existe un entorno $E(\vec{x}_0, \delta)$ de x_0 tal que para todo $\vec{x} \neq \vec{x}_0$, $\vec{x} \in E(\vec{x}_0, \delta)$ y $\vec{x} \in A$ implica que $\vec{f}(\vec{x}) \in E(\vec{l}, \epsilon)$.

La definición $\epsilon - \delta$ es una generalización directa de la definición $\epsilon - \delta$ en el caso de una variable.

Definición. Sea $\vec{f}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, donde A es un conjunto abierto. Tendremos que $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{l}$ si, y sólo si, para todo $\epsilon > 0$ existe un $\delta > 0$ tal que, si $0 < \|\vec{x} - \vec{x}_0\| < \delta$, entonces $\|\vec{f}(\vec{x}) - \vec{l}\| < \epsilon$.

Obsérvese que si el límite \vec{l} existe, el valor \vec{l} ha de ser el valor límite de $\vec{f}(\vec{x})$ a lo largo de todos los caminos hacia \vec{x}_0 . Si los distintos caminos que conducen a \vec{x}_0 no dan el mismo valor límite, el límite no existe.

Ejemplo 1. Usa la definición para verificar que se cumple el límite obvio $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{x} = \vec{x}_0$, donde \vec{x} y \vec{x}_0 están en \mathbb{R}^n .

Solución:

Sea $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ la función definida por $\vec{f}(\vec{x}) = \vec{x}$, y sea $E(\vec{x}_0, r_1)$ cualquier entorno de $\vec{x}_0 = \vec{f}(\vec{x}_0)$. Por la definición debemos hallar un entorno $E(\vec{x}_0, r_2)$ de \vec{x}_0 con la propiedad de que si $\vec{x} \neq \vec{x}_0$ y $\vec{x} \in E(\vec{x}_0, r_2)$, entonces $\vec{f}(\vec{x}) \in E(\vec{x}_0, r_1)$. Basta escoger $E(\vec{x}_0, r_2) = E(\vec{x}_0, r_1)$.

Ejemplo 2. Demuestra que no existe el límite en $(0, 0)$ de la función

$$f(x, y) = \frac{xy + y^3}{x^2 + y^2}.$$

Solución:

A lo largo de los caminos obvios hacia $(0, 0)$, los ejes coordenados, el valor límite es 0:

- si $x = 0 \implies f(x, y) = f(0, y) = y$, tiende a 0 cuando $(x, y) \rightarrow (0, 0)$;
- si $y = 0 \implies f(x, y) = f(x, 0) = 0$, tiende a 0 cuando $(x, y) \rightarrow (0, 0)$.

Sin embargo, a lo largo de la recta $y = 2x$, el valor límite es

$$f(x, y) = f(x, 2x) = \frac{2x^2 + 8x^3}{x^2 + 4x^2} = \frac{2x^2 + 8x^3}{5x^2} = \frac{2}{5} + \frac{8}{5}x \rightarrow \frac{2}{5} \quad \text{cuando } (x, y) \rightarrow (0, 0).$$

Se deduce entonces que f no tiene límite en $(0, 0)$.

Como en el caso de una variable, **el límite, si existe, es único**. Este hecho nos proporciona un **criterio** muy importante para decidir, en muchos casos, que el límite no existe.

En efecto, considerando sólo el caso de $n = 2$, si queremos calcular el $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y)$,

este límite, de existir, debe ser independiente de la forma en que (x, y) se aproxima a (x_0, y_0) . En otras palabras, si $g: \mathbb{R} \rightarrow \mathbb{R}$ es una función continua tal que $y_0 = g(x_0)$, debe suceder que si

$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l$ entonces $\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, g(x)) = l$. Este criterio nos sirve en muchos casos para probar que no existe el límite. Por ejemplo, si queremos calcular

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y^2}{x^4 + y^4}$$

y consideramos todas las rectas que pasan por el punto $(0, 0)$, $y = mx$, si existe el límite de esta función, éste debe ser el mismo que

$$\lim_{x \rightarrow 0} \frac{x^2 m^2 x^2}{x^4 + m^4 x^4}$$

10.2. LÍMITES Y CONTINUIDAD.

para cualquier valor de m . Sin embargo, en este caso es

$$\lim_{x \rightarrow 0} \frac{x^2 m^2 x^2}{x^4 + m^4 x^4} = \lim_{x \rightarrow 0} \frac{m^2 x^4}{x^4 (1 + m^4)} = \lim_{x \rightarrow 0} \frac{m^2}{1 + m^4} = \frac{m^2}{1 + m^4},$$

es decir, para valores distintos de m obtenemos distintos límites. En otras palabras, si el límite depende de los valores de m , entonces la función no tiene límite.

A esta forma de proceder se la conoce como **cálculo de límites radiales**.

Debe observarse que puede suceder que todos los límites radiales existan, sean iguales y todavía la función no tenga límite porque si nos acercamos al punto (x_0, y_0) usando otra función $y = g(x)$ que pase por (x_0, y_0) , el resultado del límite sea distinto. Por ejemplo,

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

Los límites radiales nos dan

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

Sin embargo, si escogemos $y = x^2$ resulta que

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

Como el límite de existir debe ser único, este límite no existe.

Una **propiedad** importante de los límites dobles es la siguiente:

Teorema 10.1 Si

$$\lim_{(x,y) \rightarrow (x_0, y_0)} f(x, y) = l \quad \text{entonces} \quad \left\{ \begin{array}{l} \lim_{x \rightarrow x_0} \left[\lim_{y \rightarrow y_0} f(x, y) \right] = l \\ y \\ \lim_{y \rightarrow y_0} \left[\lim_{x \rightarrow x_0} f(x, y) \right] = l. \end{array} \right.$$

A los límites $\lim_{x \rightarrow x_0} \left[\lim_{y \rightarrow y_0} f(x, y) \right]$ y $\lim_{y \rightarrow y_0} \left[\lim_{x \rightarrow x_0} f(x, y) \right]$ se les llama **límites reiterados**.

Esta propiedad nos proporciona otra **condición necesaria** para que exista el límite doble: que los límites reiterados existan y valgan lo mismo que el límite doble. Por lo tanto, si alguno de ellos no existe, o entre ellos el valor es distinto, entonces el límite doble no existe.

Ejemplo 1. Queremos saber si existe y cuánto vale

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

Solución:

Calculamos los límites reiterados

$$\lim_{x \rightarrow 0} \left[\lim_{y \rightarrow 0} \frac{x^2 - y^2}{x^2 + y^2} \right] = \lim_{x \rightarrow 0} \frac{x^2}{x^2} = \lim_{x \rightarrow 0} 1 = 1$$

y

$$\lim_{y \rightarrow 0} \left[\lim_{x \rightarrow 0} \frac{x^2 - y^2}{x^2 + y^2} \right] = \lim_{y \rightarrow 0} \frac{-y^2}{y^2} = \lim_{y \rightarrow 0} -1 = -1.$$

Ambos límites existen, pero no valen lo mismo, por consiguiente no existe el límite doble.

Otra vez, tenemos que tener cuidado: los límites reiterados pueden existir, tener el mismo valor y no existir el límite doble.

Ejemplo 2. Calcula el límite

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}.$$

Solución:

Los límites reiterados nos dan

$$\lim_{x \rightarrow 0} \left[\lim_{y \rightarrow 0} \frac{xy}{x^2 + y^2} \right] = \lim_{x \rightarrow 0} \frac{0}{x^2} = \lim_{x \rightarrow 0} 0 = 0$$

y

$$\lim_{y \rightarrow 0} \left[\lim_{x \rightarrow 0} \frac{xy}{x^2 + y^2} \right] = \lim_{y \rightarrow 0} \frac{0}{y^2} = \lim_{y \rightarrow 0} 0 = 0,$$

es decir sus dos valores coinciden.

Pero si calculamos los límites radiales: $y = mx$

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2} = \lim_{x \rightarrow 0} \frac{x(mx)}{x^2 + m^2x^2} = \lim_{x \rightarrow 0} \frac{mx^2}{x^2(1+m^2)} = \lim_{x \rightarrow 0} \frac{m}{1+m^2}.$$

Para valores distintos de m , obtenemos distintos valores del límite, por lo tanto el límite doble no existe.

Hemos visto que la existencia e igualdad de los límites radiales y reiterados es necesaria para la existencia del límite doble. Veamos ahora una **condición suficiente**.

Teorema 10.2 *Criterio de la función mayorante.*

Si existe una función $g: A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

$$i) \quad \lim_{(x,y) \rightarrow (x_0,y_0)} g(x,y) = l,$$

10.2. LÍMITES Y CONTINUIDAD.

$$ii) \quad |f(x, y) - l| \leq |g(x, y) - l|,$$

entonces

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l.$$

Demostración:

En efecto, por la hipótesis i),

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{tal que} \quad 0 < \|\vec{x} - \vec{x}_0\| < \delta \quad \Rightarrow \quad |g(\vec{x}) - l| < \epsilon.$$

Y por la hipótesis ii),

$$|f(x, y) - l| \leq |g(x, y) - l| < \epsilon.$$

Así que

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l.$$

c.q.d

Ejemplo. Calcula

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{\sqrt{x^2 + y^2}}.$$

Solución:

Empecemos calculando cuánto valen los límites radiales. $y = mx$

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{\sqrt{x^2 + y^2}} = \lim_{x \rightarrow 0} \frac{x(mx)}{\sqrt{x^2 + m^2 x^2}} = \lim_{x \rightarrow 0} \frac{mx^2}{\sqrt{x^2(1 + m^2)}} = \lim_{x \rightarrow 0} \frac{mx}{\sqrt{1 + m^2}} = 0.$$

Luego, de existir el límite, éste debe ser cero.

Veamos que existe una función $g(x, y)$ tal que

$$\lim_{(x,y) \rightarrow (x_0,y_0)} g(x, y) = 0 \quad \text{y} \quad |f(x, y) - 0| \leq |g(x, y) - 0|.$$

En efecto,

$$|f(x, y) - 0| = \left| \frac{xy}{\sqrt{x^2 + y^2}} \right| \leq \left| \frac{xy}{\sqrt{x^2}} \right| \leq |y|.$$

En otras palabras, poniendo $g(x, y) = y$, se cumplen las dos condiciones del teorema 10.2.

Hay otro criterio, a veces útil, que consiste en pasar a coordenadas polares. Utilizarlo correctamente no siempre es sencillo, así que no lo estudiaremos.

Por otro lado, en algunas ocasiones el límite doble puede reducirse a un límite de una función de una variable.

Teorema 10.3 Supongamos que $f: A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ sea una función que se puede poner como composición de dos funciones $f = h \circ g$, siendo $g: A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ y $h: \mathbb{R} \rightarrow \mathbb{R}$. Supongamos, además, que existe el

$$\lim_{(x,y) \rightarrow (x_0,y_0)} g(x,y) = t_0.$$

Si

$$\lim_{t \rightarrow t_0} h(t) = l$$

entonces

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = l.$$

Demostración:

En efecto, si $\lim_{t \rightarrow t_0} h(t) = l$,

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \text{tal que} \quad 0 < |t - t_0| < \delta \quad \Rightarrow \quad |h(t) - l| < \epsilon.$$

Para este $\delta > 0$, como $\lim_{(x,y) \rightarrow (x_0,y_0)} g(x,y) = t_0$,

$$\exists \delta_1 > 0 \quad \text{tal que} \quad 0 < \|(x,y) - (x_0,y_0)\| < \delta_1 \quad \Rightarrow \quad |g(x,y) - t_0| < \delta.$$

Por lo anterior

$$|f(x,y) - l| \leq |h(g(x,y)) - l| < \epsilon \quad \Leftrightarrow \quad |f(x,y) - l| < \epsilon.$$

Así que

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x,y) = l.$$

c.q.d

Ejemplo. Calcula

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2 + \text{sen}(x+y)}{x+y}.$$

Solución:

En primer lugar notemos que

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2 + \text{sen}(x+y)}{x+y} = \lim_{(x,y) \rightarrow (0,0)} \left[(x-y) + \frac{\text{sen}(x+y)}{x+y} \right] = \lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen}(x+y)}{x+y}.$$

Sea entonces

$$f(x,y) = \frac{\text{sen}(x+y)}{x+y}, \quad f = h \circ g \quad \text{con} \quad g(x,y) = x+y \quad \text{y} \quad h(t) = \frac{\text{sen } t}{t}.$$

Así pues, como

$$\lim_{(x,y) \rightarrow (0,0)} g(x,y) = \lim_{(x,y) \rightarrow (0,0)} x+y = 0 \quad \text{y} \quad \lim_{t \rightarrow 0} h(t) = \lim_{t \rightarrow 0} \frac{\text{sen } t}{t} = 1$$

concluimos que

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 - y^2 + \text{sen}(x+y)}{x+y} = \lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen}(x+y)}{x+y} = 1.$$

10.2. LÍMITES Y CONTINUIDAD.

Para realizar cálculos prácticos necesitamos algunas **reglas**.

Sean $\vec{f}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\vec{g}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, con \vec{x}_0 un elemento de A o un punto de frontera de A . Además, si $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{l}_1$ y $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{g}(\vec{x}) = \vec{l}_2$, entonces:

- $\lim_{\vec{x} \rightarrow \vec{x}_0} [\vec{f}(\vec{x}) + \vec{g}(\vec{x})] = \vec{l}_1 + \vec{l}_2$;
- $\lim_{\vec{x} \rightarrow \vec{x}_0} [\alpha \vec{f}(\vec{x})] = \alpha \vec{l}_1$, donde $\alpha \in \mathbb{R}$;
- si $m = 1$, $\lim_{\vec{x} \rightarrow \vec{x}_0} [f(\vec{x}) \cdot g(\vec{x})] = l_1 \cdot l_2$;
- si $m = 1$, $\lim_{\vec{x} \rightarrow \vec{x}_0} \left[\frac{f(\vec{x})}{g(\vec{x})} \right] = \frac{l_1}{l_2}$, con $l_2 \neq 0$;
- si $\vec{f} = (f_1, f_2, \dots, f_m)$ con $f_i: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ las funciones componentes de \vec{f} :
 $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{l} = (l_1, l_2, \dots, l_m)$ si, y sólo si, $\lim_{\vec{x} \rightarrow \vec{x}_0} f_i(\vec{x}) = l_i$.

Ejemplo. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = x^3 + y^2 - xy + 1$. Calcula

$$\lim_{(x,y) \rightarrow (0,1)} f(x, y).$$

Solución:

Aquí f es la suma de las cuatro funciones $(x, y) \rightsquigarrow x^3$, $(x, y) \rightsquigarrow y^2$, $(x, y) \rightsquigarrow -xy$, $(x, y) \rightsquigarrow 1$. El límite de una suma es la suma de los límites, y el límite de un producto es el producto de los límites. Usando el hecho de que $\lim_{(x,y) \rightarrow (x_0, y_0)} x = x_0$ y $\lim_{(x,y) \rightarrow (x_0, y_0)} y = y_0$, obtenemos

$$\lim_{(x,y) \rightarrow (0,1)} f(x, y) = 0^3 + 1^2 - 0 \cdot 1 + 1 = 2.$$

En el estudio de funciones de una variable, aprendimos que el concepto de función continua está basado en la idea intuitiva de una función cuya gráfica es una curva sin romper. Esto es, una curva sin saltos, o el tipo de curva que se trazaría al mover la punta de un lápiz sin separarla del papel. La misma idea sirve para funciones de varias variables. Formulamos estos conceptos en términos de límites.

Definición. Sean $\vec{f}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\vec{x}_0 \in A$. Decimos que \vec{f} es **continua en** \vec{x}_0 si y sólo si

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{f}(\vec{x}_0),$$

que es equivalente a decir

$$\lim_{\vec{h} \rightarrow \vec{0}} \vec{f}(\vec{x} + \vec{h}) = \vec{f}(\vec{x}_0).$$

Si decimos simplemente que \vec{f} es **continua**, queremos decir que \vec{f} es continua en cada punto \vec{x}_0 de A .

Como la condición $\lim_{\vec{x} \rightarrow \vec{x}_0} \vec{f}(\vec{x}) = \vec{f}(\vec{x}_0)$ significa que $\vec{f}(\vec{x})$ está cerca de $\vec{f}(\vec{x}_0)$ cuando \vec{x} está cerca de \vec{x}_0 , vemos que nuestra definición corresponde al requerimiento de que la gráfica de \vec{f} no esté rota.

Ejemplo. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por $f(x, y) = xy$. Demuestra que f es continua.

Solución:

Usando el hecho de que $\lim_{(x,y) \rightarrow (x_0, y_0)} x = x_0$ y $\lim_{(x,y) \rightarrow (x_0, y_0)} y = y_0$, y por las propiedades de los límites, obtenemos

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = \lim_{(x,y) \rightarrow (x_0, y_0)} (xy) = \left(\lim_{(x,y) \rightarrow (x_0, y_0)} x \right) \left(\lim_{(x,y) \rightarrow (x_0, y_0)} y \right) = x_0 y_0 = f(\vec{x}_0).$$

Se puede ver, por el mismo método, que cualquier polinomio de varias variables es continuo en todo punto. En el caso de dos variables, esto significa la continuidad en todo punto del plano xy ; en el caso de tres variables, esto significa la continuidad en todo el espacio. Las funciones racionales (los cocientes de polinomios) son continuas en todos los puntos exceptos aquellos en los cuales se anula el denominador, que son los puntos que no pertenecen al dominio de existencia de la función.

Ejemplos. Las siguientes funciones son continuas en el dominio indicado a lado:

- $f(x, y) = \frac{2x - y}{x^2 + y^2}$ continua en $\{(x, y) \in \mathbb{R}^2 \mid (x, y) \neq (0, 0)\}$ (excepto el origen);
- $f(x, y) = \frac{1}{x^2 - y}$ continua en $\{(x, y) \in \mathbb{R}^2 \mid y \neq x^2\}$ (excepto la parábola $y = x^2$);
- $F(x, y, z) = \frac{x^4 - y^3}{x - z}$ continua en $\{(x, y, z) \in \mathbb{R}^3 \mid x \neq z\}$ (excepto el plano $x = z$).

Una función continua de varias variables es continua con respecto a cada una de sus variables por separado. El recíproco es falso. Es posible que una función sea continua en cada una de sus variables separadamente sin ser continua como función de varias variables.

Ejemplo. Determina si es continua en $(0, 0)$ la función

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

Solución:

Dado que $f(x, 0) = 0$ para todo x y $f(0, y) = 0$ para todo y tenemos que en el punto $(0, 0)$, la función f es continua en las variables x e y . Sin embargo, si examinamos los valores que toma la función f en los puntos situados sobre la recta $y = kx$, con $k = \text{const.}$, tenemos

$$f(x, kx) = \frac{2xkx}{x^2 + k^2x^2} = \frac{2k}{1 + k^2},$$

es decir, sobre cada recta que pasa por el origen de coordenadas, la función tiene un valor constante que depende del coeficiente angular k de la recta. Por eso el límite de la función depende del camino que recorra, lo que significa que la función no tiene límite. Por consiguiente, la función es discontinua en este punto. Para ser continua f tendría que tender a $f(0, 0) = 0$. Además, está claro que no se puede modificar el valor de esta función en el origen de coordenadas para convertirla en continua.

10.2. LÍMITES Y CONTINUIDAD.

Las reglas dadas para calcular los límites permiten deducir las siguientes **propiedades** de la continuidad. Sean $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $g: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, con \bar{x}_0 un elemento de A . Entonces:

- si \vec{f} y \vec{g} son continuas en \bar{x}_0 también lo es $\vec{f} + \vec{g}$, donde $(\vec{f} + \vec{g})(\bar{x}) = \vec{f}(\bar{x}) + \vec{g}(\bar{x})$;
- si \vec{f} es continua en \bar{x}_0 , también lo es $\alpha \vec{f}$, donde $\alpha \vec{f}(\bar{x}) = \alpha [\vec{f}(\bar{x})]$;
- si $m = 1$, \vec{f} y \vec{g} son continuas en \bar{x}_0 también lo es $f \cdot g$, donde $(f \cdot g)(\bar{x}) = f(\bar{x}) \cdot g(\bar{x})$;
- si $m = 1$, f es continua en \bar{x}_0 y no se anula en A , también lo es $\frac{1}{f}$, donde $\left(\frac{1}{f}\right)(\bar{x}) = \frac{1}{f(\bar{x})}$;
- si $\vec{f} = (f_1, f_2, \dots, f_m)$ con $f_i: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ las funciones componentes de \vec{f} , entonces \vec{f} es continua en \bar{x}_0 si y sólo si cada una de las componentes f_i es continua en \bar{x}_0 .

Ejemplo. Sea $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por

$$\vec{f}(x, y) = \left(x^2 + y^2, \frac{x^2 + y^2}{1 + x^2} \right).$$

Muestra que \vec{f} es continua.

Solución:

Para verlo, es suficiente mostrar que cada componente es continua. La primera componente es continua por ser suma de dos funciones continuas $(x, y) \rightsquigarrow x^2$ y $(x, y) \rightsquigarrow y^2$. La segunda componente también es continua por ser producto de la función continua $x^2 + y^2$ por la función $1/(1 + x^2)$ que es continua porque la función $1 + x^2$ es continua y no nula.

Estudiemos ahora la composición, otra operación básica que puede efectuarse con funciones.

Definición. Si \vec{g} está definida de $A \subset \mathbb{R}^n$ a $B \subset \mathbb{R}^m$, y \vec{f} está definida de $B \subset \mathbb{R}^m$ a $C \subset \mathbb{R}^p$, la **composición de \vec{g} con \vec{f}** , o de \vec{f} en \vec{g} , denotada por $\vec{f} \circ \vec{g}$ manda A a C mediante $\bar{x} \rightsquigarrow \vec{f}(\vec{g}(\bar{x}))$.

Por la composición de funciones tenemos la siguiente **propiedad**:

Sean $\vec{g}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\vec{f}: B \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$. Sea además $\vec{g}(A) \subset B$, de manera que $\vec{f} \circ \vec{g}$ está definida en A . Si \vec{g} es continua en $\bar{x}_0 \in A$ y \vec{f} es continua en $\vec{y}_0 = \vec{g}(\bar{x}_0)$, entonces la composición $\vec{f} \circ \vec{g}$ es continua en \bar{x}_0 .

Ejemplo. Muestra que la función $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, $f(x, y, z) = (x^2 + y^2 + z^2)^3 + \text{sen}(xyz)$, es continua.

Solución:

Podemos escribir f como suma de las dos funciones $f_1: (x, y, z) \rightsquigarrow (x^2 + y^2 + z^2)^3$ y $f_2: (x, y, z) \rightsquigarrow \text{sen}(xyz)$, de modo que basta mostrar que cada una es continua. La primera es composición de la función continua $(x, y, z) \rightsquigarrow x^2 + y^2 + z^2$ con la función también continua $u \rightsquigarrow u^3$, y la segunda es la composición de $(x, y, z) \rightsquigarrow xyz$ con $u \rightsquigarrow \text{sen } u$, que son ambas continuas.

Ejercicios.

- 1) En cada uno de los casos siguientes, sea S el conjunto de todos los puntos (x, y) del plano o (x, y, z) del espacio tri-dimensional que satisfacen las desigualdades dadas. Determina si S es o no abierto. En los casos de conjunto en el plano, dibuja un gráfico mostrando el conjunto S , explica geoméricamente si es abierto o no, e indica la frontera de S en el gráfico.

- | | |
|--|---|
| a) $x^2 + y^2 < 1$; | b) $3x^2 + 2y^2 < 6$; |
| c) $x^2 + y^2 \leq 1$; | d) $1 < x^2 + y^2 < 2$; |
| e) $ x < 1$ y $ y < 1$; | f) $x > y$; |
| g) $ x \leq 1$ y $ y \leq 1$; | h) $x \geq y$; |
| i) $ x \leq 1, y < 1$ y $ z < 1$; | j) $x + y + z < 1, x > 0, y > 0, z > 0$. |

- 2) Calcula los siguientes límites.

- | | |
|--|--|
| a) $\lim_{(x,y) \rightarrow (0,1)} x^3 y$; | b) $\lim_{(x,y) \rightarrow (0,1)} e^x y$; |
| c) $\lim_{(x,y) \rightarrow (0,1)} \frac{\operatorname{sen}^2 x}{x^2}$; | d) $\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2 + 3)$; |
| e) $\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2 + 2}$; | f) $\lim_{(x,y) \rightarrow (0,0)} \frac{e^x}{x + 1}$; |
| g) $\lim_{(x,y) \rightarrow (0,0)} \frac{\cos x - 1 - x^2/2}{x^4 + y^4}$; | h) $\lim_{(x,y) \rightarrow (0,0)} \frac{(x - y)^2}{x^2 + y^2}$; |
| i) $\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$; | j) $\lim_{(x,y) \rightarrow (0,0)} \frac{xy^2}{(x^2 + y^2)^{3/2}}$. |

- 3) Sea la función $\bar{f}: \mathbb{R} \rightarrow \mathbb{R}^2$, dada por $x \rightsquigarrow (x^2, e^x)$, calcula $\lim_{x \rightarrow 1} \bar{f}(x)$.

- 4) Muestra que $\lim_{(x,y) \rightarrow (0,0)} \frac{\operatorname{sen}(xy)}{xy} = 1$.

- 5) Demuestra que las funciones siguientes no tienen límite en $(0, 0)$.

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

- 6) Determina si la función $f(x, y) = \frac{x - y^4}{x^3 - y^4}$ tiene límite en $(1, 1)$.

10.3. DIFERENCIACIÓN.

$$7) \text{ Sea } f(x, y) = \begin{cases} x \operatorname{sen} \frac{1}{y} & y \neq 0 \\ 0 & y = 0. \end{cases}$$

Demuestra que $f(x, y) \rightarrow 0$ cuando $(x, y) \rightarrow (0, 0)$ pero que

$$\lim_{y \rightarrow 0} \left[\lim_{x \rightarrow 0} f(x, y) \right] \neq \lim_{x \rightarrow 0} \left[\lim_{y \rightarrow 0} f(x, y) \right].$$

8) Si $(x, y) \neq (0, 0)$, sea $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$. Halla los límites de $f(x, y)$ cuando $(x, y) \rightarrow (0, 0)$ a lo largo de la recta $y = mx$. ¿Es posible definir $f(0, 0)$ de modo que f sea continua en $(0, 0)$?

9) Determina para cada función el conjunto de puntos (x, y) en los que f es continua.

a) $f(x, y) = x^4 + y^4 - 4x^2y^2;$

b) $f(x, y) = \operatorname{tg} \frac{x^2}{y};$

c) $f(x, y) = \frac{x}{\sqrt{x^2 + y^2}};$

d) $f(x, y) = \operatorname{arctg} \frac{x + y}{1 - xy}.$

10.3 Diferenciación.

Con los métodos vistos hasta ahora para dibujar gráficas de funciones de varias variables, a veces puede ser imposible obtener suficiente información para comprender incluso las características generales de una función complicada. En el caso de funciones reales de una variable real, sabemos que el concepto de derivada puede ser de mucha ayuda: por ejemplo, nos permite localizar máximos y mínimos. Ya hemos comentado que una función continua no tiene la gráfica rota. Una función diferenciable de \mathbb{R}^2 a \mathbb{R} debe ser tal que su gráfica no esté rota, pero además debe tener bien definido un plano tangente a la gráfica en cada punto. Así, no debe haber dobleces, esquinas o picos en la gráfica. Es decir, la gráfica debe ser suave. Para una función de una variable, la diferenciabilidad garantiza la continuidad (ver teorema 1.1). Como veremos, para las funciones de varias variables, la existencia de las derivadas parciales no basta para garantizar la continuidad.

Intuitivamente: la existencia de una derivada parcial sólo depende del comportamiento de la función a lo largo de un segmento de recta, mientras que la continuidad depende del comportamiento de la función en todas las direcciones.

Para precisar estas ideas necesitamos algunas definiciones.

Definición. Sea $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una función con valores reales, con U un conjunto abierto. Las **derivadas parciales de f** con respecto a la primera, segunda, ..., n -ésima variable son las funciones con valores reales, de n variables, definidas por

$$\begin{aligned} \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) &= \lim_{h \rightarrow 0} \frac{f(x_1, x_2, \dots, x_j + h, \dots, x_n) - f(x_1, \dots, x_n)}{h} = \\ &= \lim_{h \rightarrow 0} \frac{f(\vec{x} + h \vec{e}_j) - f(\vec{x})}{h} = f_{x_j} \end{aligned}$$

si existen los límites, donde $1 \leq j \leq n$ y \vec{e}_j es el j -ésimo vector de la base usual.

Claramente la derivada parcial de f con respecto a la variable x_j , f_{x_j} , es la derivada de la función f con respecto a la variable x_j , manteniendo las otras $(n - 1)$ variables fijas. Si f es función de tres variables, con frecuencia usaremos la notación $f_x = \partial f / \partial x$, $f_y = \partial f / \partial y$, $f_z = \partial f / \partial z$, en lugar de $\partial f / \partial x_1$, $\partial f / \partial x_2$, $\partial f / \partial x_3$. Si $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, tenemos $\vec{f}(x_1, x_2, \dots, x_n) = (f_1(x_1, x_2, \dots, x_n), \dots, f_m(x_1, x_2, \dots, x_n))$, de modo que podemos hablar de las derivadas parciales de cada componente; por ejemplo $\partial f_m / \partial x_n$ es la derivada parcial de la m -ésima componente de \vec{f} con respecto a x_n , la n -ésima variable.

Para indicar que una derivada parcial ha de evaluarse en algún punto en particular, por ejemplo para denotar la derivada parcial con respecto a x en el punto (x_0, y_0) escribimos

$$f_x(x_0, y_0) \quad \circ \quad \frac{\partial f}{\partial x}(x_0, y_0) \quad \circ \quad \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)}$$

En el caso de una variable, sabemos que $f'(x_0)$ proporciona el coeficiente de variación instantánea con respecto a x de $f(x)$ en $x = x_0$.

En el caso de dos variables, $f_x(x_0, y_0)$ da el coeficiente de variación con respecto a x de $f(x, y_0)$ en $x = x_0$ y $f_y(x_0, y_0)$ da el coeficiente de variación con respecto a y de $f(x_0, y)$ en $y = y_0$.

A estos números se le puede dar **una interpretación geométrica**. Si representamos la superficie $z = f(x, y)$ y consideramos la sección de la superficie con el plano $y = y_0$, que es un plano paralelo al plano xz , obtenemos que esta sección es la gráfica de la función $g(x) = f(x, y_0)$. Diferenciando con respecto a x , tenemos $g'(x) = f_x(x, y_0)$ y en $x = x_0$ será $g'(x_0) = f_x(x_0, y_0)$. Por consiguiente, el número $f_x(x_0, y_0)$ es la pendiente de la sección en $y = y_0$ de la superficie $z = f(x, y)$ en el punto $P(x_0, y_0, f(x_0, y_0))$. De manera análoga para la otra derivada parcial, tenemos que el número $f_y(x_0, y_0)$ es la pendiente de la sección en $x = x_0$ de la superficie $z = f(x, y)$ en el punto $P(x_0, y_0, f(x_0, y_0))$.

En el caso de tres variables, $f_x(x_0, y_0, z_0)$ da el coeficiente de variación con respecto a x de $f(x, y_0, z_0)$ en $x = x_0$, $f_y(x_0, y_0, z_0)$ da el coeficiente de variación con respecto a y de $f(x_0, y, z_0)$ en $y = y_0$ y $f_z(x_0, y_0, z_0)$ da el coeficiente de variación con respecto a z de $f(x_0, y_0, z)$ en $z = z_0$. Sin embargo, el significado geométrico que hemos dado a las derivadas parciales de una función de dos variables no tiene significado análogo en el caso de tres o más variables.

Ejemplo 1. Para la función

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

tenemos

$$f_x(x, y) = \frac{\partial f}{\partial x}(x, y) = y e^{xy} + \frac{2x}{x^2 + y} \quad \text{y} \quad f_y(x, y) = \frac{\partial f}{\partial y}(x, y) = x e^{xy} + \frac{1}{x^2 + y}.$$

En la primera derivada parcial con respecto a x hemos mantenido la y constante y hemos diferenciado sólo con respecto a x ; de la misma manera con la derivada parcial con respecto

10.3. DIFERENCIACIÓN.

a y hemos mantenido la x constante y hemos diferenciado sólo con respecto a y . El número $f_x(2, 1) = e^2 + \frac{2}{5}$ da el coeficiente de variación con respecto a x de la función $f(x, 1) = e^x + \ln(x^2 + 1)$ en $x = 2$; así mismo el número $f_y(2, 1) = 2e^2 + \frac{1}{5}$ da el coeficiente de variación con respecto a y de la función $f(2, y) = e^{2y} + \ln(4 + y)$ en $y = 1$.

Ejemplo 2. Para la función

$$f(x, y, z) = x y^2 z^3$$

las derivadas parciales son

$$f_x(x, y, z) = y^2 z^3, \quad f_y(x, y, z) = 2 x y z^3, \quad f_z(x, y, z) = 3 x y^2 z^2.$$

Ejemplo 3. Para una función de la forma

$$f(x, y, z) = F(x, y) G(y, z)$$

podemos escribir

$$f_x(x, y, z) = F_x(x, y) G(y, z), \quad f_z(x, y, z) = F(x, y) G_z(y, z),$$

$$f_y(x, y, z) = F_y(x, y) G(y, z) + F(x, y) G_y(y, z).$$

Vamos a ver ahora un ejemplo que nos aclare como puede ser insuficiente una definición de diferenciabilidad que requiera sólo de la existencia de las derivadas parciales. En este caso no se cumple la regla de la cadena para varias variables.

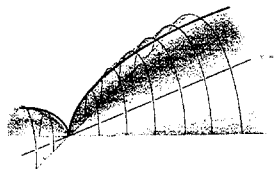
Sea $f(x, y) = \sqrt[3]{x} \sqrt[3]{y}$. Esta función está definida en todo \mathbb{R}^2 . Por definición,

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

y, de manera similar

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0.$$

Es necesario usar la definición original de derivadas parciales, por el hecho de que las funciones $\sqrt[3]{x}$ y $\sqrt[3]{y}$ no son diferenciables en 0. Consideremos ahora la composición, $(f \circ g)$, de la función $g: \mathbb{R} \rightarrow \mathbb{R}^2$, definida por $g(x) = (x, x)$, con $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por $f(x, y) = \sqrt[3]{x} \sqrt[3]{y}$. Resulta que la composición está dada por $(f \circ g)(x) = \sqrt[3]{x^2}$. Cada componente de la función g es diferenciable en x , y la función f tiene derivadas parciales en $(0, 0)$, pero $(f \circ g)$ no es diferenciable en $x = 0$, en el sentido del cálculo de una variable.



En contraste con lo que sucede con el cálculo de funciones de una variable, donde la composición de funciones diferenciables es diferenciable, en el caso de funciones de varias variables la composición de f y g no es diferenciable. Para rectificar esta situación tendremos que dar una definición de diferenciabilidad que nos permita obtener que la composición de funciones diferenciables sea diferenciable. Hay otro motivo para no estar satisfecho con la sola existencia de las derivadas parciales de $f(x, y) = \sqrt[3]{x} \sqrt[3]{y}$: no hay plano tangente a la gráfica en $(0, 0)$. Dado que $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, si hubiera plano tangente, debería ser el plano xy . Sin embargo, como resulta evidente de la figura, la gráfica tiene una severa arruga en el punto $(0, 0)$.

Para motivar la definición de diferenciabilidad, calculemos cuál tendría que ser la ecuación del plano tangente a la gráfica de $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ en el punto genérico (x_0, y_0) . En \mathbb{R}^3 , un plano no vertical tiene una ecuación de la forma

$$z = ax + by + c$$

Si se trata de que sea el plano tangente a la gráfica de f , las pendientes a lo largo de los ejes x e y deben ser iguales a $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$, que son las velocidades de cambio de f respecto a x e y . Así,

$$a = \frac{\partial f}{\partial x}(x_0, y_0) \quad \text{y} \quad b = \frac{\partial f}{\partial y}(x_0, y_0).$$

Finalmente podemos determinar la constante c a partir del hecho de que $z = f(x_0, y_0)$, cuando $x = x_0, y = y_0$:

$$\begin{aligned} f(x_0, y_0) &= \frac{\partial f}{\partial x}(x_0, y_0) x_0 + \frac{\partial f}{\partial y}(x_0, y_0) y_0 + c \implies \\ c &= f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0) x_0 - \frac{\partial f}{\partial y}(x_0, y_0) y_0. \end{aligned}$$

Luego, obtenemos

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

ecuación del plano tangente a la gráfica de f en (x_0, y_0) , si f es suficientemente suave. Nuestra definición de diferenciabilidad dirá que el plano dado por la ecuación así obtenida es una buena aproximación de f cerca del punto (x_0, y_0) .

Definición. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. Decimos que f es **diferenciable en** (x_0, y_0) si las derivadas parciales $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ existen en (x_0, y_0) y si

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

10.3. DIFERENCIACIÓN.

Si denotamos con $\mathbf{D}f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$ al vector fila de las derivadas parciales, la definición de diferenciabilidad afirma que

$$f(x_0, y_0) + \mathbf{D}f(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

es una buena aproximación de la función f .

Antes hemos usado la idea intuitiva, ahora podemos definir el plano tangente a la gráfica.

Definición. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ diferenciable en (x_0, y_0) . El plano en \mathbb{R}^3 definido por

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \quad (10.1)$$

se llama **plano tangente** a la gráfica de f en el punto (x_0, y_0) .

Ejemplo. Calcula el plano tangente a la gráfica de $z = x^3 + y^3 + e^{xy}$ en el punto $(1, 0, 2)$.

Solución:

Usamos la fórmula (10.1), con $x_0 = 1$, $y_0 = 0$, $z_0 = f(1, 0) = 2$. Las derivadas parciales son

$$\frac{\partial f}{\partial x} = 3x^2 + ye^{xy}, \quad \frac{\partial f}{\partial y} = 3y^2 + xe^{xy} \quad \implies \quad \frac{\partial f}{\partial x}(1, 0) = 3, \quad \frac{\partial f}{\partial y}(1, 0) = 1,$$

luego, sustituyendo en la fórmula (10.1)

$$z = 2 + 3(x - 1) + 1(y - 0) \quad \implies \quad z = 3x + y - 1.$$

Podemos ahora pasar a las funciones \vec{f} de \mathbb{R}^n a \mathbb{R}^m . La derivada $\mathbf{D}\vec{f}(\vec{x}_0)$ de la función $\vec{f} = (f_1, \dots, f_m)$ en un punto $\vec{x}_0 = (x_1, \dots, x_n)$ es ahora una matriz $m \times n$ con elementos $t_{ij} = \frac{\partial f_i}{\partial x_j}$, con $i = 1, \dots, m$, $j = 1, \dots, n$.

$\mathbf{D}\vec{f}(\vec{x}_0)$ se llama **matriz de las derivadas parciales de \vec{f} en \vec{x}_0** , comúnmente llamada **matriz jacobiana**.

Definición. Sean U un conjunto abierto en \mathbb{R}^n y $\vec{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Decimos que \vec{f} es **diferenciable en $\vec{x}_0 \in U$** si existen las derivadas parciales de \vec{f} en \vec{x}_0 (es decir si está bien definida la matriz de las derivadas parciales $\mathbf{D}\vec{f}(\vec{x}_0)$) y si

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{\|\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0) - \mathbf{D}\vec{f}(\vec{x}_0)(\vec{x} - \vec{x}_0)\|}{\|\vec{x} - \vec{x}_0\|} = 0.$$

Ejemplo. Calcula la matriz de las derivadas parciales para la función $\vec{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $\vec{f}(x, y, z) = (e^{x+y}, xyz)$.

Solución:

$$D\vec{f}(\vec{x}) = D\vec{f}(x, y, z) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{pmatrix} = \begin{pmatrix} e^{x+y} & e^{x+y} & 0 \\ yz & xz & xy \end{pmatrix}.$$

Definición. Si consideramos el caso especial $m = 1$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$, la matriz de las derivadas $Df(\vec{x})$ es una matriz $1 \times n$, es decir está formada por solo una fila.

$$Df(\vec{x}) = \left(\frac{\partial f}{\partial x_1}(\vec{x}) \quad \frac{\partial f}{\partial x_2}(\vec{x}) \quad \dots \quad \frac{\partial f}{\partial x_n}(\vec{x}) \right).$$

Al nuevo vector así conseguido

$$\left(\frac{\partial f}{\partial x_1}(\vec{x}), \frac{\partial f}{\partial x_2}(\vec{x}), \dots, \frac{\partial f}{\partial x_n}(\vec{x}) \right) = \vec{\nabla} f(\vec{x}) = \text{grad } f(\vec{x})$$

se le llama **gradiente de f** , y viene denotado por $\text{grad } f$ ó $\vec{\nabla} f$.

De la definición vemos que para $f: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$\vec{\nabla} f(\vec{x}) = \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j},$$

y para $f: \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$\vec{\nabla} f(\vec{x}) = \left(\frac{\partial f}{\partial x}(x, y, z), \frac{\partial f}{\partial y}(x, y, z), \frac{\partial f}{\partial z}(x, y, z) \right) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k}.$$

El significado geométrico lo daremos más adelante.

Ejemplo. Calcula el vector gradiente para la función $f(x, y, z) = z e^{2x+3y} + x y z$.

Solución:

Podemos escribir el vector gradiente en las dos formas

$$\vec{\nabla} f(\vec{x}) = \left((2z e^{2x+3y} + yz), (3z e^{2x+3y} + xz), (e^{2x+3y} + xy) \right)$$

$$\vec{\nabla} f(\vec{x}) = (2z e^{2x+3y} + yz) \vec{i} + (3z e^{2x+3y} + xz) \vec{j} + (e^{2x+3y} + xy) \vec{k}.$$

Para una función de una variable, sabemos que la diferenciabilidad garantiza la continuidad. Más adelante enunciaremos que esto también se cumple para las funciones diferenciables de varias variables. Sin embargo, para las funciones de varias variables, la sola existencia de las derivadas parciales no basta para garantizar la continuidad. Veamos los siguientes ejemplos. El primero considera una de las funciones cuya continuidad hemos estudiado en la sección 10.2.

10.3. DIFERENCIACIÓN.

Ejemplo 1. Ya sabemos que la función

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

no es continua en $(0, 0)$, sin embargo las dos derivadas parciales existen en el origen, dado que $f(x, 0)$ y $f(0, y)$ valen siempre cero:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{(2h \cdot 0)/(h^2 + 0) - 0}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0,$$

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{(2 \cdot 0 \cdot h)/(0 + h^2) - 0}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0.$$

Ejemplo 2. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} 1 & \text{si } x = 0 \text{ o si } y = 0 \\ 0 & \text{de no ser así.} \end{cases}$$

Como f es constante en los ejes x e y , donde es igual a 1, tenemos que las dos derivadas parciales existen en el origen:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{1 - 1}{h} = 0,$$

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{1 - 1}{h} = 0.$$

Pero f no es continua en $(0, 0)$, pues $\lim_{(x, y) \rightarrow (0, 0)} f(x, y)$ no existe. Para ver este último resultado, podemos examinar los valores que toma la función f en los puntos situados sobre la recta $y = kx$, con $k = \text{const.}$, tenemos $f(x, kx) = 0$, si $x \neq 0$. Pero si nos ponemos en uno de los ejes, por ejemplo en el eje y , con ecuación $x = 0$, tenemos $f(0, y) = 1$. No coincidiendo estos valores, el límite no existe. Por consiguiente, la función es discontinua en este punto. Para ser continua f tendría que tender a $f(0, 0) = 1$.

No es difícil entender cómo una función puede tener derivadas parciales y no ser continuas, si pensamos en el hecho de que la existencia de una derivada parcial depende del comportamiento de la función a lo largo de un segmento de recta, mientras que la continuidad depende del comportamiento de la función en todas las direcciones.

Enunciaremos ahora dos teoremas de los cuales no daremos la demostración por ser algo complicada. El lector interesado puede encontrarla en uno cualquiera de los textos indicados en la bibliografía.

Teorema 10.4 Sea $\vec{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en $\vec{x}_0 \in U$. Entonces \vec{f} es continua en \vec{x}_0 .

Como ya hemos visto, por lo general es fácil decir si existen las derivadas parciales usando los conocimientos del cálculo de funciones de una variable. Sin embargo, resulta mucho más complicado usar la definición de función diferenciable, y comprobar que

$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{\|\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0) - \mathbf{D}\vec{f}(\vec{x}_0)(\vec{x} - \vec{x}_0)\|}{\|\vec{x} - \vec{x}_0\|} = 0$. Afortunadamente existe un criterio sencillo, que nos dice cuándo una función es diferenciable.

Teorema 10.5 Sea $\vec{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$. Supongamos que existen todas las derivadas parciales $\frac{\partial f_i}{\partial x_j}$ de \vec{f} y son continuas en un entorno de un punto $\vec{x} \in U$. Entonces \vec{f} es diferenciable en \vec{x} .

Definición. Una función cuyas derivadas parciales existan y sean continuas, se dice de **clase \mathcal{C}^1** .

Así, el teorema 10.5 dice que cualquier función \mathcal{C}^1 es diferenciable. Es interesante constatar la siguiente cadena de implicaciones:

$$\text{Parciales continuas} \implies \text{Diferenciable} \implies \text{Existen las parciales}$$

la primera implicación es por el teorema 10.5 y la segunda se sigue de la definición de derivada. Nótese que todos los enunciados recíprocos obtenidos invirtiendo una implicación no son válidos. Un contraejemplo al recíproco de la primera implicación viene dado por la función

$$f(x) = \begin{cases} x^2 \operatorname{sen}(1/x) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

Para esta función tenemos

$$\frac{df}{dx}(x) = 2x \operatorname{sen}\left(\frac{1}{x}\right) + x^2 \cos\left(\frac{1}{x}\right) \left(-\frac{1}{x^2}\right) = 2x \operatorname{sen}\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

Como contraejemplo al recíproco de la segunda podemos considerar la función

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

Nótese que esta función es continua: por el criterio de la función mayorante podemos escribir, para todo $(x, y) \neq (0, 0)$

$$\left| \frac{xy}{\sqrt{x^2 + y^2}} \right| \leq \left| \frac{xy}{\sqrt{x^2}} \right| \leq |y|,$$

que tiende hacia cero cuando $(x, y) \rightarrow (0, 0)$; luego, $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = f(0, 0) = 0$.

Además,

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{(h \cdot 0) / \sqrt{h^2 + 0} - 0}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0,$$

10.3. DIFERENCIACIÓN.

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h \rightarrow 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h \rightarrow 0} \frac{(0h)/\sqrt{0+h^2} - 0}{h} = \lim_{h \rightarrow 0} \frac{0-0}{h} = 0.$$

Así, las derivadas parciales existen en $(0,0)$. Además, si $(x,y) \neq (0,0)$, tenemos

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

que no tiene límite cuando $(x,y) \rightarrow (0,0)$: en los puntos situados sobre la recta $y = kx$, con $k = \text{const.}$, tenemos

$$\frac{\partial f}{\partial x}(x, kx) = \frac{kx}{\sqrt{x^2+k^2x^2}} - \frac{x^2 kx}{(x^2+k^2x^2)^{3/2}} = \frac{k}{\sqrt{1+k^2}} - \frac{k}{(1+k^2)^{3/2}},$$

es decir se obtienen límites diferentes. Luego las derivadas parciales no son continuas en $(0,0)$ y por lo tanto no podemos usar el teorema 10.5. Podemos ahora intentar demostrar que f no es diferenciable, sin embargo, como hemos visto, es continua. Aplicamos la definición de diferenciability: decir que

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{\|\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0) - \mathbf{D}\vec{f}(\vec{x}_0)(\vec{x} - \vec{x}_0)\|}{\|\vec{x} - \vec{x}_0\|} = 0$$

quiere decir que para cualquier $\epsilon > 0$ existiría un $\delta > 0$ tal que para $0 < \|(h_1, h_2)\| < \delta$ implique que

$$\frac{\left\| f(h_1, h_2) - f(0,0) - (0,0) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|}{\|(h_1, h_2)\|} < \epsilon$$

esto es,

$$\frac{\|f(h_1, h_2) - f(0,0)\|}{\|(h_1, h_2)\|} < \epsilon \iff \frac{|h_1 h_2|}{h_1^2 + h_2^2} < \epsilon \iff |h_1 h_2| < \epsilon(h_1^2 + h_2^2).$$

Pero si escogemos $h_1 = h_2$, queda que $\frac{1}{2} < \epsilon$, que no sería cierto si escogiéramos $\epsilon \leq \frac{1}{2}$. Por lo tanto, f no es diferenciable en $(0,0)$. Si se hace una gráfica de la función notaremos que la función está torcida en $(0,0)$.

Ejercicios.

1) Halla las derivadas parciales de primer orden de las siguientes funciones:

a) $z = \log(x + \sqrt{x^2 + y^2}),$

b) $z = x^y,$

c) $u = (x y^7),$

d) $z = \arctg \frac{y}{x};$

e) $z = \sqrt{x^2 + y^2},$

f) $z = \frac{x+y}{x-y}, \quad x \neq y.$

g) $z = x^2 + y^2 \text{ sen}(xy),$

h) $z = \frac{x}{\sqrt{x^2 + y^2}}, \quad (x,y) \neq (0,0);$

- i) $z = x^4 + y^4 - 4x^2y^2$, j) $z = \log(x^2 + y^2)$, $(x, y) \neq (0, 0)$;
 k) $z = \frac{1}{y} \cos x^2$, l) $z = x^{(y^2)}$, $x > 0$.

2) Halla el vector gradiente $\vec{\nabla}f(x, y)$ en cada punto de \mathbb{R}^2 , en el que exista, para cada una de las siguientes funciones:

- a) $f(x, y) = x^2 + y^2 \operatorname{sen}(xy)$;
 b) $f(x, y) = x^2 y^3 z^4$;
 c) $f(x, y) = x^2 y^2 \log(x^2 + y^2)$, si $(x, y) \neq (0, 0)$ y $f(0, 0) = 0$;
 d) $f(x, y) = xy \operatorname{sen}\left(\frac{1}{x^2 + y^2}\right)$, si $(x, y) \neq (0, 0)$ y $f(0, 0) = 0$.

3) Demuestra que $x \frac{\partial z}{\partial x} + y \frac{\partial z}{\partial y} = xy + z$, si $z(x, y) = xy + x e^{(y/x)}$.

4) Demuestra que $\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} + \frac{\partial u}{\partial z} = 1$, si $u(x, y, z) = x + \frac{x-y}{y-z}$.

5) Sea $v(r, t) = t^n \exp\left(\frac{-r^2}{4t}\right)$. Halla un valor de la constante n tal que la función $v(r, t)$ satisfaga la ecuación $\frac{\partial v}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial v}{\partial r} \right)$.

10.4 Propiedades de la derivada.

En el cálculo elemental aprendimos cómo diferenciar sumas, productos, cocientes y funciones compuestas. Ahora queremos generalizar estas ideas a funciones de varias variables, con particular atención a la diferenciación de funciones compuestas. En el caso de una variable, la regla para diferenciación de composiciones, llamada **regla de la cadena**, nos da la siguiente fórmula

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx} = f'(g(x)) g'(x)$$

donde f es una función de una variable real a valores reales, $z = f(y)$, e y es una función de x , $y = g(x)$, es decir $z = f(g(x))$. Antes de dar la fórmula equivalente al caso de varias variables, comencemos con las reglas de diferenciación para sumas, productos y cocientes.

(1) **Regla del múltiplo constante.** Sea $\vec{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciable en \vec{x}_0 y sea c un número real. Entonces $\vec{h}(\vec{x}) = c\vec{f}(\vec{x})$ es diferenciable en \vec{x}_0 y

$$\mathbf{D}\vec{h}(\vec{x}_0) = c\mathbf{D}\vec{f}(\vec{x}_0);$$

10.4. PROPIEDADES DE LA DERIVADA.

- (2) **Regla de la suma.** Sean $\vec{f}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\vec{g}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ diferenciables en \vec{x}_0 . Entonces $\vec{h}(\vec{x}) = \vec{f}(\vec{x}) + \vec{g}(\vec{x})$ es diferenciable en \vec{x}_0 y

$$D\vec{h}(\vec{x}_0) = D\vec{f}(\vec{x}_0) + D\vec{g}(\vec{x}_0);$$

- (3) **Regla del producto.** Sean $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciables en \vec{x}_0 . Entonces $h(\vec{x}) = f(\vec{x}) \cdot g(\vec{x})$ es diferenciable en \vec{x}_0 y

$$Dh(\vec{x}_0) = g(\vec{x}_0) Df(\vec{x}_0) + f(\vec{x}_0) Dg(\vec{x}_0);$$

- (4) **Regla del cociente.** Sean $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciables en \vec{x}_0 , con g nunca nula en U . Entonces $h(\vec{x}) = f(\vec{x})/g(\vec{x})$ es diferenciable en \vec{x}_0 y

$$Dh(\vec{x}_0) = \frac{g(\vec{x}_0) Df(\vec{x}_0) - f(\vec{x}_0) Dg(\vec{x}_0)}{[g(\vec{x}_0)]^2};$$

Nótese que en los últimos dos casos $m = 1$, es decir las últimas dos reglas son válidas solo para funciones de varias variables a valores reales.

Ejemplo. Verifica la fórmula para Dh en la regla (4) con $f(x, y, z) = 2x^3 + y^3 + 3z^3$ y $g(x, y, z) = y^2 + 1$.

Solución:

Aquí, tenemos

$$h(x, y, z) = \frac{f(x, y, z)}{g(x, y, z)} = \frac{2x^3 + y^3 + 3z^3}{y^2 + 1}.$$

Por diferenciación directa tenemos que las derivadas parciales son

$$\frac{\partial h}{\partial x} = \frac{6x^2}{y^2 + 1}, \quad \frac{\partial h}{\partial z} = \frac{9z^2}{y^2 + 1},$$

$$\frac{\partial h}{\partial y} = \frac{3y^2(y^2 + 1) - (2x^3 + y^3 + 3z^3)2y}{(y^2 + 1)^2} = \frac{y(y^3 + 3y - 4x^3 - 6z^3)}{(y^2 + 1)^2}$$

entonces

$$Dh(x, y, z) = \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z} \right) = \left(\frac{6x^2}{y^2 + 1}, \frac{y(y^3 + 3y - 4x^3 - 6z^3)}{(y^2 + 1)^2}, \frac{9z^2}{y^2 + 1} \right).$$

Por la regla (4), obtenemos

$$\begin{aligned} Dh(\vec{x}) &= \frac{g(\vec{x}) Df(\vec{x}) - f(\vec{x}) Dg(\vec{x})}{[g(\vec{x})]^2} = \\ &= \frac{(y^2 + 1) (6x^2, 3y^2, 9z^2) - (2x^3 + y^3 + 3z^3) (0, 2y, 0)}{(y^2 + 1)^2} \end{aligned}$$

que es lo mismo que se ha obtenido directamente.

Pasamos ahora a enunciar el teorema que nos da la **regla de la cadena** para funciones de varias variables, que si usamos la notación matricial D se parece a la regla para una variable.

Teorema 10.6 Regla de la cadena.

Sean $U \subset \mathbb{R}^n$ y $V \subset \mathbb{R}^m$ dos conjuntos abiertos. Sean, además, $\vec{g}: U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $\vec{f}: V \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$ dos funciones dadas tales que $g(U) \subset V$, de modo que está definida la composición $\vec{f} \circ \vec{g}$. Si \vec{g} es diferenciable en \vec{x}_0 y \vec{f} es diferenciable en $\vec{y}_0 = \vec{g}(\vec{x}_0)$, entonces $\vec{f} \circ \vec{g}$ es diferenciable en \vec{x}_0 y

$$D(\vec{f} \circ \vec{g})(\vec{x}_0) = D\vec{f}(\vec{y}_0) \cdot D\vec{g}(\vec{x}_0).$$

No daremos la demostración de este teorema, pero vamos a considerar dos casos particulares que son importantes por sí mismos.

Primer caso especial de la regla de la cadena. Supongamos que $\vec{c}: \mathbb{R} \rightarrow \mathbb{R}^3$ y $f: \mathbb{R}^3 \rightarrow \mathbb{R}$. Sea $h(t) = f(\vec{c}(t)) = f(x(t), y(t), z(t))$, con $\vec{c}(t) = (x(t), y(t), z(t))$. Entonces

$$\frac{dh}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} + \frac{\partial f}{\partial z} \frac{dz}{dt}.$$

Esto es,

$$\frac{dh}{dt} = \vec{\nabla} f(\vec{c}(t)) \cdot \vec{c}'(t).$$

Éste es el caso especial del teorema 10.6 en el que $\vec{c} = \vec{g}$, f con valores reales, $n = p = 1$ y $m = 3$. Nótese que

$$\vec{\nabla} f(\vec{c}(t)) \cdot \vec{c}'(t) = Df(\vec{c}(t)) \cdot D\vec{c}(t),$$

donde el producto en el lado izquierdo es el producto escalar de vectores, mientras que el producto del lado derecho es la multiplicación de matrices: $Df(\vec{c}(t))$ es una matriz fila y $D\vec{c}(t)$ es una matriz columna. Los vectores $\vec{\nabla} f(\vec{c}(t))$ y $\vec{c}'(t)$ tienen las mismas componentes que sus equivalentes matriciales. El cambio de notación indica el paso de matrices a vectores. Como veremos más adelante (ver sección 10.7), la función $\vec{c}: \mathbb{R} \rightarrow \mathbb{R}^3$ describe una trayectoria en \mathbb{R}^3 .

Segundo caso especial de la regla de la cadena. Supongamos que $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ y $\vec{g}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, con $\vec{g}(x, y, z) = (u(x, y, z), v(x, y, z), w(x, y, z))$. Sea $h: \mathbb{R}^3 \rightarrow \mathbb{R}$ dada por $h(x, y, z) = f(\vec{g}(x, y, z)) = f(u(x, y, z), v(x, y, z), w(x, y, z))$. Entonces

$$\left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z} \right) = \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w} \right) \cdot \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix}.$$

En este caso especial del teorema 10.6 hemos considerado $n = m = 3$ y $p = 1$. Nótese que se ha escrito explícitamente el producto de matrices. Si realizamos la multiplicación matricial obtenemos las tres ecuaciones

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial x},$$

10.4. PROPIEDADES DE LA DERIVADA.

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial y},$$

$$\frac{\partial h}{\partial z} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial z} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial z} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial z}.$$

Está claro cómo se puede generalizar a valores distintos de n , m y p . Como un ejemplo más consideremos el caso de $f: \mathbb{R}^4 \rightarrow \mathbb{R}$ y $\vec{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^4$, con \vec{g} dada por $\vec{g}(x, y) = (u(x, y), v(x, y), w(x, y), z(x, y))$. Luego, si $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ viene dada por $h(x, y) = f(\vec{g}(x, y)) = f(u(x, y), v(x, y), w(x, y), z(x, y))$, tenemos

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial x} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial x}$$

y

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} + \frac{\partial f}{\partial w} \frac{\partial w}{\partial y} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial y}.$$

Ejemplo 1. Halla du/dt si

$$u = x^2 - y^2 \quad y \quad x = t^2 - 1, \quad y = 3 \operatorname{sen} \pi t.$$

Solución:

Estamos en el caso de dos variables x e y , que dependen de una nueva variable t , luego

$$\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt}.$$

Dado que

$$\frac{\partial u}{\partial x} = 2x, \quad \frac{\partial u}{\partial y} = -2y \quad y \quad \frac{dx}{dt} = 2t, \quad \frac{dy}{dt} = 3\pi \cos \pi t$$

tenemos

$$\frac{du}{dt} = (2x)(2t) - (2y)(3\pi \cos \pi t) = 4xt - 6y\pi \cos \pi t$$

y substituyendo las expresiones de x e y :

$$\frac{du}{dt} = 4t^3 - 4t - 18\pi \operatorname{sen} \pi t \cos \pi t.$$

Se obtiene el mismo resultado escribiendo primero u directamente como función de t

$$u = x^2 - y^2 = (t^2 - 1)^2 - (3 \operatorname{sen} \pi t)^2$$

y diferenciando después

$$\frac{du}{dt} = 2(t^2 - 1)(2t) - 2(3 \operatorname{sen} \pi t)(3\pi \cos \pi t) = 4t^3 - 4t - 18\pi \operatorname{sen} \pi t \cos \pi t.$$

Ejemplo 2. Verifica la regla de la cadena para

$$f(u, v, w) = u^3 - v^2 + w$$

donde

$$u(x, y, z) = x^2 y, \quad v(x, y, z) = y^2 z, \quad w(x, y, z) = xy e^z.$$

Solución:

Aquí, estamos con un ejemplo del segundo caso especial de la regla de la cadena:

$$h(x, y, z) = f(u(x, y, z), v(x, y, z), w(x, y, z)) = x^6 y^3 - y^4 z^2 + xy e^z.$$

Diferenciando directamente obtenemos las tres ecuaciones

$$\frac{\partial h}{\partial x} = 6x^5 y^3 + y e^z, \quad \frac{\partial h}{\partial y} = 3x^6 y^2 - 4y^3 z^2 + x e^z, \quad \frac{\partial h}{\partial z} = -2y^4 z + xy e^z.$$

Por otro lado, usando la regla de la cadena, las dos matrices serán

$$\begin{aligned} \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w} \right) &= (3u^2, -2v, 1) = (3x^4 y^2, -2y^2 z, 1) \\ \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix} &= \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2yz & y^2 \\ ye^z & xe^z & xye^z \end{pmatrix}. \end{aligned}$$

Realizando la multiplicación llegamos a los mismos resultados:

$$\begin{aligned} \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z} \right) &= (3x^4 y^2, -2y^2 z, 1) \cdot \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2yz & y^2 \\ ye^z & xe^z & xye^z \end{pmatrix} = \\ &= (6x^5 y^3 + ye^z, 3x^6 y^2 - 4y^3 z^2 + xe^z, -2y^4 z + xye^z). \end{aligned}$$

Ejemplo 3. Sean $\vec{f}(x, y) = (\cos y + x^3, e^{x+y} + y^2)$ y $\vec{g}(u, v) = (e^{uv}, u - \sin v)$.

(a) Escribir una fórmula para $\vec{f} \circ \vec{g}$;

(b) Calcular $\mathbf{D}(\vec{f} \circ \vec{g})(0, 0)$.

Solución:

En este caso tenemos $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y $\vec{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, luego $\vec{f} \circ \vec{g}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Para la primera parte tenemos

$$\vec{f} \circ \vec{g} = f(e^{uv}, u - \sin v) = (\cos(u - \sin v) + e^{3uv}, \exp(e^{uv} + u - \sin v) + (u - \sin v)^2).$$

Para la segunda parte, recordemos que $\mathbf{D}(\vec{f} \circ \vec{g})$ será una matriz 2×2 . Sin duda será conveniente usar la regla de la cadena, en lugar de realizar la diferenciación de las expresiones obtenidas arriba. Por la regla de la cadena,

$$\mathbf{D}(\vec{f} \circ \vec{g})(0, 0) = \mathbf{D}(\vec{f}(\vec{g}(0, 0))) \cdot \mathbf{D}(\vec{g}(0, 0)) = \mathbf{D}(\vec{f}(1, 0)) \cdot \mathbf{D}(\vec{g}(0, 0)).$$

10.4. PROPIEDADES DE LA DERIVADA.

Ahora

$$D(\vec{f}(1,0)) = \left(\begin{array}{cc} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{array} \right) \Bigg|_{(x,y)=(1,0)} = \left(\begin{array}{cc} 3x^2 & -\operatorname{sen} y \\ e^{x+y} & e^{x+y} + 2y \end{array} \right) \Bigg|_{(x,y)=(1,0)} = \left(\begin{array}{cc} 3 & 0 \\ e & e \end{array} \right),$$

$$D(\vec{g}(0,0)) = \left(\begin{array}{cc} \frac{\partial g_1}{\partial u} & \frac{\partial g_1}{\partial v} \\ \frac{\partial g_2}{\partial u} & \frac{\partial g_2}{\partial v} \end{array} \right) \Bigg|_{(u,v)=(0,0)} = \left(\begin{array}{cc} v e^{uv} & u e^{uv} \\ 1 & -\cos v \end{array} \right) \Bigg|_{(u,v)=(0,0)} = \left(\begin{array}{cc} 0 & 0 \\ 1 & -1 \end{array} \right).$$

Así,

$$D(\vec{f} \circ \vec{g})(0,0) = \left(\begin{array}{cc} 3 & 0 \\ e & e \end{array} \right) \cdot \left(\begin{array}{cc} 0 & 0 \\ 1 & -1 \end{array} \right) = \left(\begin{array}{cc} 0 & 0 \\ e & -e \end{array} \right).$$

Ejercicios.

- 1) Sean las funciones $f(x,y) = x^2 + y$, $h(u) = (\operatorname{sen} 3u, \operatorname{cos} 8u)$. Sea $g(u) = f(h(u))$.
Calcula $\frac{dg}{du}$ en $u = 0$ de ambas formas, directamente y usando la regla de la cadena.
- 2) Demuestra que la función $w = f(u,v)$, donde $u = x + at$, $v = y + bt$ verifica la ecuación $\frac{\partial w}{\partial t} = a \frac{\partial w}{\partial x} + b \frac{\partial w}{\partial y}$.
- 3) Las ecuaciones $u = f(x,y)$, $x = X(t)$, $y = Y(t)$ definen u como función de t , pongamos $u = F(t)$. Calcula $F'(t)$ en función de t en cada uno de los siguientes casos particulares.
 - a) $f(x,y) = x^2 + y^2$, $X(t) = t$, $Y(t) = t^2$;
 - b) $f(x,y) = e^{xy} \operatorname{cos}(xy^2)$, $X(t) = \operatorname{cos} t$, $Y(t) = \operatorname{sen} t$;
 - c) $f(x,y) = \log \left(\frac{1 + e^{x^2}}{1 + e^{y^2}} \right)$, $X(t) = e^t$, $Y(t) = e^{-t}$.
- 4) Comprueba la regla de la cadena para $\frac{\partial h}{\partial x}$ donde $h(x,y) = f(u(x,y), v(x,y))$ y $f(u,v) = \frac{u^2 + v^2}{u^2 - v^2}$, con $u(x,y) = e^{-x-y}$, $v(x,y) = e^{xy}$.
- 5) Sean $\vec{f}(u,v) = (\operatorname{tg}(u-1) - e^v, u^2 - v^2)$ y $\vec{g}(x,y) = (e^{x-y}, x-y)$. Calcula $\vec{f} \circ \vec{g}$ y $D(\vec{f} \circ \vec{g})(1,1)$.
- 6) Verifica el primer caso especial de la regla de la cadena para la composición $f \circ \vec{c}$ en cada uno de los casos siguientes.
 - a) $f(x,y) = xy$, $\vec{c}(t) = (e^t, \operatorname{cos} t)$;

- b) $f(x, y) = e^{xy}$, $\vec{c}(t) = (3t^2, t^3)$;
 c) $f(x, y) = (x^2 + y^2) \log \sqrt{x^2 + y^2}$, $\vec{c}(t) = (e^t, e^{-t})$;
 d) $f(x, y) = x \exp(x^2 + y^2)$, $\vec{c}(t) = (t, -t)$.

10.5 Gradientes y derivadas direccionales.

En la sección 10.1 hemos estudiado la gráfica de las funciones de varias variables con valores reales y en la sección 10.3 hemos introducido la definición de gradiente. Ahora queremos estudiar las funciones de varias variables con valores reales usando los métodos del cálculo y usar gradientes para obtener una fórmula para el plano tangente a una superficie de nivel. En esta sección nos restringimos al caso de funciones de tres variables reales a valores reales. Pero, antes de todo, recordemos cómo se define el gradiente.

Definición. Si $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ es diferenciable, el **gradiente de f** en (x, y, z) es el vector en el espacio \mathbb{R}^3 dado por

$$\text{grad } f = \vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right).$$

Así, $\vec{\nabla} f$ es simplemente la matriz de las derivadas Df , escrita como vector.

Ejemplo 1. Para $f(x, y, z) = \text{sen}(xy^2z^3)$ tenemos

$$\frac{\partial f}{\partial x} = y^2 z^3 \cos(xy^2z^3), \quad \frac{\partial f}{\partial y} = 2xy z^3 \cos(xy^2z^3), \quad \frac{\partial f}{\partial z} = 3xy^2 z^2 \cos(xy^2z^3).$$

Luego el vector gradiente será

$$\vec{\nabla} f = (y^2 z^3 \cos(xy^2z^3), 2xy z^3 \cos(xy^2z^3), 3xy^2 z^2 \cos(xy^2z^3)).$$

que podemos también escribir como

$$\vec{\nabla} f = \cos(xy^2z^3) [y^2 z^3 \vec{i} + 2xy z^3 \vec{j} + 3xy^2 z^2 \vec{k}].$$

Ejemplo 2. Sea $f(x, y, z) = x \text{sen}(\pi y) + y \cos(\pi z)$. Calcula $\vec{\nabla} f$ en el punto $(0, 1, 2)$.

Solución:

Empezamos hallando el vector gradiente en un punto genérico (x, y, z) :

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = (\text{sen}(\pi y), \pi x \cos(\pi y) + \cos(\pi z), -\pi y \text{sen}(\pi z)).$$

Y evaluando en $(0, 1, 2)$ obtenemos

$$\vec{\nabla} f(0, 1, 2) = (\text{sen } \pi, \pi(0) \cos \pi + \cos(2\pi), -\pi(1) \text{sen}(2\pi)) = (0, 1, 0) \iff \vec{\nabla} f(0, 1, 2) = \vec{j}.$$

10.5. GRADIENTES Y DERIVADAS DIRECCIONALES.

Ejemplo 3. Sea $f(x, y, z) = \sqrt{x^2 + y^2 + z^2} = r$, la distancia del origen al punto (x, y, z) . Si denotamos el punto con el vector posición $\vec{r} = (x, y, z)$ tenemos que

$$\vec{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) = \frac{\vec{r}}{r}.$$

Es decir, $\vec{\nabla} f$ es el vector unitario en la dirección de $\vec{r} = (x, y, z)$.

Ejemplo 4. Sea $f(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{1}{r}$. Entonces

$$\vec{\nabla} f = \left(\frac{-x}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-y}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-z}{(x^2 + y^2 + z^2)^{3/2}} \right) = -\frac{\vec{r}}{r^3}.$$

En algunos aspectos, el gradiente se comporta como las derivadas en el caso de una variable. En particular, si $\vec{\nabla} f(\vec{x})$ y $\vec{\nabla} g(\vec{x})$ existen, también existen $\vec{\nabla} [f(\vec{x}) + g(\vec{x})]$, $\vec{\nabla} [\alpha f(\vec{x})]$ y $\vec{\nabla} [f(\vec{x}) \cdot g(\vec{x})]$, y como se puede fácilmente comprobar se verifica que

$$\vec{\nabla} [f(\vec{x}) + g(\vec{x})] = \vec{\nabla} f(\vec{x}) + \vec{\nabla} g(\vec{x}), \quad \vec{\nabla} [\alpha f(\vec{x})] = \alpha \vec{\nabla} f(\vec{x}),$$

$$\vec{\nabla} [f(\vec{x}) \cdot g(\vec{x})] = f(\vec{x}) \vec{\nabla} g(\vec{x}) + g(\vec{x}) \vec{\nabla} f(\vec{x}).$$

Pasamos ahora a considerar las derivadas direccionales. Supongamos que f sea una función a valores reales, es decir, $f: \mathbb{R}^3 \rightarrow \mathbb{R}$. Sean \vec{v} y $\vec{x} \in \mathbb{R}^3$ unos vectores fijos, podemos considerar entonces la función de \mathbb{R} a \mathbb{R} definida por $t \rightarrow f(\vec{x} + t\vec{v})$. El conjunto de puntos de la forma $\vec{x} + t\vec{v}$, con $t \in \mathbb{R}$, es la recta L que pasa por el punto \vec{x} y paralela al vector \vec{v} . Por lo tanto, la función $f(\vec{x} + t\vec{v})$ representa la función f restringida a la recta L . Podemos preguntarnos: ¿con qué rapidez están cambiando los valores de f a lo largo de la recta L en el punto \vec{x} ? Como la razón de cambio de una función está dada por una derivada, podemos responder que es el valor de la derivada de esta función de t en $t = 0$, dado que cuando $t = 0$, $\vec{x} + t\vec{v}$ se reduce a \vec{x} . Podemos formalizar este concepto con la siguiente definición.

Definición. Si $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, la derivada direccional de f en \vec{x} en la dirección de un vector \vec{v} está dada por

$$D_{\vec{v}} f(\vec{x}) = \left. \frac{d}{dt} f(\vec{x} + t\vec{v}) \right|_{t=0}.$$

Esto es, si existe el límite

$$\lim_{h \rightarrow 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h}.$$

Se llama también tasa de cambio de f en la dirección del vector \vec{v} .

Teorema 10.7 Si $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ es diferenciable, entonces existen todas las derivadas direccionales. La derivada direccional en \vec{x} en la dirección $\vec{v} = (v_1, v_2, v_3)$ está dada por

$$D_{\vec{v}}f(\vec{x}) = \vec{\nabla} f(\vec{x}) \cdot \vec{v} = v_1 \frac{\partial f}{\partial x} + v_2 \frac{\partial f}{\partial y} + v_3 \frac{\partial f}{\partial z}.$$

Demostración:

Sea $\vec{c}(t) = \vec{x} + t\vec{v}$, de manera que $f(\vec{x} + t\vec{v}) = f(\vec{c}(t))$. Además, $\vec{c}(0) = \vec{x}$ y $\vec{c}'(0) = \vec{v}$. Por el primer caso especial de la regla de la cadena, se sigue que

$$\frac{d}{dt}f(\vec{c}(t)) = \vec{\nabla} f(\vec{c}(t)) \cdot \vec{c}'(t)$$

y evaluando en $t = 0$ tenemos

$$\left. \frac{d}{dt}f(\vec{x} + t\vec{v}) \right|_{t=0} = \vec{\nabla} f(\vec{c}(0)) \cdot \vec{c}'(0) = \vec{\nabla} f(\vec{x}) \cdot \vec{v}.$$

c.q.d.

En la definición de derivada direccional, es conveniente escoger \vec{v} como un vector unitario, para que el resultado no dependa de la longitud del vector considerado. Nótese que si se considera $\alpha\vec{v}$ como nuevo vector con la misma dirección de \vec{v} , tenemos $D_{\alpha\vec{v}}f(\vec{x}) = \vec{\nabla} f(\vec{x}) \cdot \alpha\vec{v} = \alpha D_{\vec{v}}f(\vec{x})$.

Está claro que las propias derivadas parciales son derivadas direccionales:

$$\frac{\partial f}{\partial x} = D_{\vec{i}}f(\vec{x}) = \vec{\nabla} f(\vec{x}) \cdot \vec{i} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \cdot (1, 0, 0)$$

así mismo

$$\frac{\partial f}{\partial y} = D_{\vec{j}}f(\vec{x}) \quad \text{y} \quad \frac{\partial f}{\partial z} = D_{\vec{k}}f(\vec{x}).$$

Ejemplo. Sea $f(x, y, z) = x \cos y \sin z$. Calcular la tasa de cambio de f en la dirección del vector $\vec{v} = 2\vec{i} - \vec{j} + 4\vec{k}$ en el punto $(1, \pi, \pi/4)$.

Solución:

Dado que $\|\vec{v}\| = \sqrt{4 + 1 + 16} = \sqrt{21}$, el vector unitario en la dirección de $2\vec{i} - \vec{j} + 4\vec{k}$ es el vector

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\sqrt{21}}{21} (2\vec{i} - \vec{j} + 4\vec{k}).$$

Además,

$$\frac{\partial f}{\partial x} = \cos y \sin z, \quad \frac{\partial f}{\partial y} = -x \sin y \sin z, \quad \frac{\partial f}{\partial z} = x \cos y \cos z.$$

10.5. GRADIENTES Y DERIVADAS DIRECCIONALES.

de manera que

$$\frac{\partial f}{\partial x}(1, \pi, \pi/4) = -\frac{\sqrt{2}}{2}, \quad \frac{\partial f}{\partial y}(1, \pi, \pi/4) = 0, \quad \frac{\partial f}{\partial z}(1, \pi, \pi/4) = -\frac{\sqrt{2}}{2}.$$

Luego,

$$\vec{\nabla} f(1, \pi, \pi/4) = -\frac{\sqrt{2}}{2}\vec{i} - \frac{\sqrt{2}}{2}\vec{k} = -\frac{\sqrt{2}}{2}[\vec{i} + \vec{k}]$$

y por lo tanto

$$D_{\vec{u}}f(1, \pi, \pi/4) = \vec{\nabla} f(1, \pi, \pi/4) \cdot \vec{u} = \left(\frac{\sqrt{21}}{21}\right) \left(-\frac{\sqrt{2}}{2}\right) (2 + 4) = -\frac{\sqrt{42}}{7}.$$

Por la definición de proyección dada en la sección 9.2 es fácil identificar la derivada direccional en el sentido de \vec{u} como la componente del gradiente de f en esa dirección.

De aquí, podemos dar una interpretación geométrica del gradiente:

Teorema 10.8 *Supongamos que $\vec{\nabla} f \neq \vec{0}$. Entonces, $\vec{\nabla} f$ apunta en la dirección a lo largo de la cual f crece más rápido.*

Demostración:

Sea \vec{u} un vector unitario, entonces por el teorema 10.7 la tasa de cambio de f en la dirección de \vec{u} es $D_{\vec{u}}f(\vec{x}) = \vec{\nabla} f(\vec{x}) \cdot \vec{u} = \|\vec{\nabla} f(\vec{x})\| \|\vec{u}\| \cos \theta = \|\vec{\nabla} f(\vec{x})\| \cos \theta$, donde θ es el ángulo entre $\vec{\nabla} f(\vec{x})$ y \vec{u} . Dado que $-1 \leq \cos \theta \leq 1$, tenemos que para todas las direcciones \vec{u} , $-\|\vec{\nabla} f(\vec{x})\| \leq D_{\vec{u}}f(\vec{x}) \leq \|\vec{\nabla} f(\vec{x})\|$. Si \vec{u} apunta en el sentido de $\vec{\nabla} f(\vec{x})$, entonces la derivada direccional es máxima, es decir el coeficiente de variación vale $\|\vec{\nabla} f(\vec{x})\|$, y al contrario es mínima si apunta en el sentido contrario, es decir el coeficiente de variación vale $-\|\vec{\nabla} f(\vec{x})\|$.

c.q.d.

En otras palabras, si queremos movernos en una dirección en la cual f va creciendo más rápidamente, debemos proceder en la dirección de $\vec{\nabla} f(\vec{x})$. Análogamente, si deseamos movernos en una dirección en la cual f decrece más rápido, deberemos proceder en la dirección de $-\vec{\nabla} f(\vec{x})$.

Ejemplo. La densidad (masa por unidad de volumen) de una bola de metal centrada en el origen viene dada por la función

$$\lambda(x, y, z) = k \exp[-(x^2 + y^2 + z^2)],$$

con k constante positiva.

- ¿En qué dirección crece la densidad más rápidamente a partir del punto (x, y, z) ? ¿Cuál es el coeficiente de variación del incremento de densidad?
- ¿En qué dirección decrece más rápidamente la densidad?

- c) ¿Cuáles son los coeficientes de variación de la densidad en (x, y, z) en las direcciones de los tres ejes coordenados?

Solución:

El gradiente para esta función será el vector

$$\begin{aligned}\bar{\nabla} \lambda(x, y, z) &= \frac{\partial \lambda}{\partial x}(x, y, z) \bar{i} + \frac{\partial \lambda}{\partial y}(x, y, z) \bar{j} + \frac{\partial \lambda}{\partial z}(x, y, z) \bar{k} \\ &= -2k \exp[-(x^2 + y^2 + z^2)] [x \bar{i} + y \bar{j} + z \bar{k}] = \\ &= -2\lambda(x, y, z) \bar{r},\end{aligned}$$

que claramente apunta desde (x, y, z) en sentido opuesto al radio vector. Entonces

- a) La densidad crece más rápidamente hacia el origen.

El coeficiente de variación del crecimiento de densidad es

$$\|\bar{\nabla} \lambda(x, y, z)\| = 2\lambda(x, y, z) \|\bar{r}\| = 2\lambda(x, y, z) \sqrt{x^2 + y^2 + z^2}.$$

- b) La densidad decrece más rápidamente alejándose desde el origen.

- c) Los coeficientes de variación de la densidad en (x, y, z) en las direcciones de los tres ejes coordenados \bar{i} , \bar{j} y \bar{k} vienen dados por las derivadas direccionales

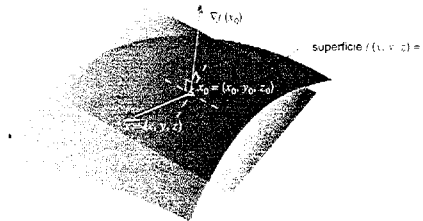
$$D_{\bar{i}}\lambda(x, y, z) = \bar{\nabla} \lambda(x, y, z) \cdot \bar{i} = -2x\lambda(x, y, z),$$

$$D_{\bar{j}}\lambda(x, y, z) = \bar{\nabla} \lambda(x, y, z) \cdot \bar{j} = -2y\lambda(x, y, z),$$

$$D_{\bar{k}}\lambda(x, y, z) = \bar{\nabla} \lambda(x, y, z) \cdot \bar{k} = -2z\lambda(x, y, z),$$

que coinciden con las derivadas parciales de λ .

Ahora veremos la relación entre el gradiente de una función f y sus superficies de nivel. El gradiente apunta en la dirección en la que los valores de f cambian más rápidamente, mientras que una superficie de nivel está en la dirección en la que esos valores no cambian. Si f es suficientemente suave, el gradiente y la superficie de nivel serán perpendiculares.



Teorema 10.9 Sean $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ una función de clase C^1 y (x_0, y_0, z_0) un punto en la superficie de nivel S definida por $f(x_0, y_0, z_0) = k$, para k constante. Entonces $\bar{\nabla} f(x_0, y_0, z_0)$ es normal a la superficie de nivel en el sentido siguiente: si \bar{v} es el vector tangente en $t = 0$ de una trayectoria $\bar{c}(t)$ en S , con $\bar{c}(0) = (x_0, y_0, z_0)$, entonces $\bar{\nabla} f \cdot \bar{v} = 0$.

10.5. GRADIENTES Y DERIVADAS DIRECCIONALES.

Demostración:

Sea $\vec{c}(t)$ en S ; entonces $f(\vec{c}(t)) = k$. Sea \vec{v} como en la hipótesis; entonces $\vec{v} = \vec{c}'(0)$. Así, el hecho de que $f(\vec{c}(t))$ es constante en t y la regla de la cadena dan

$$0 = \left. \frac{d}{dt} f(\vec{c}(t)) \right|_{t=0} = \vec{\nabla} f(\vec{c}(0)) \cdot \vec{c}'(0) = \vec{\nabla} f(\vec{x}) \cdot \vec{v}.$$

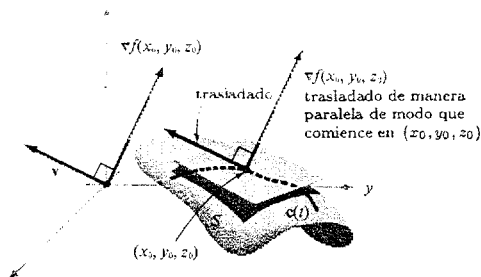
c.q.d.

Es entonces razonable dar la siguiente definición.

Definición. Sea S la superficie formada por los puntos (x, y, z) tales que $f(x, y, z) = k$, para k constante. El **plano tangente** de S en un punto (x_0, y_0, z_0) de S está definido por la ecuación

$$\vec{\nabla} f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0, \quad (10.2)$$

$$\text{si } \vec{\nabla} f(x_0, y_0, z_0) = \frac{\partial f}{\partial x}(x_0, y_0, z_0) \vec{i} + \frac{\partial f}{\partial y}(x_0, y_0, z_0) \vec{j} + \frac{\partial f}{\partial z}(x_0, y_0, z_0) \vec{k} \neq \vec{0}.$$



Esta es una ecuación para el plano tangente en notación vectorial. En coordenadas cartesianas, la ecuación toma la forma

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

Esto extiende la definición que dimos antes, ver fórmula (10.1), para el plano tangente a la gráfica de una función.

En el caso de una función de dos variables, la superficie de nivel se reduce a $f(x, y) = k$, para k constante, que como ya sabemos es una curva de nivel, C . El vector gradiente será entonces un vector, en el plano xy , perpendicular a C en (x_0, y_0) , que llamaremos **vector normal**, $\vec{\nabla} f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) \vec{i} + \frac{\partial f}{\partial y}(x_0, y_0) \vec{j}$.

La recta que pasa por (x_0, y_0) y es perpendicular al gradiente es la **recta tangente**, cuya ecuación en notación vectorial resulta ser

$$\vec{\nabla} f(x_0, y_0) \cdot (x - x_0, y - y_0) = 0.$$

En coordenadas cartesianas, la ecuación toma la forma

$$(x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$



Por otro lado, el vector $\vec{t} = \frac{\partial f}{\partial y}(x_0, y_0) \vec{i} - \frac{\partial f}{\partial x}(x_0, y_0) \vec{j}$ resulta ser perpendicular al gradiente, es decir es un **vector tangente** a la curva. La recta que pasa por (x_0, y_0) y es perpendicular al vector tangente es la **recta normal**, cuya ecuación en notación vectorial resulta ser

$$\vec{t}(x_0, y_0) \cdot (x - x_0, y - y_0) = 0.$$

En coordenadas cartesianas, la ecuación toma la forma

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

Nótese que una superficie dada en la forma $z = g(x, y)$ puede reescribirse en la forma $f(x, y, z) = 0$, definiendo $f(x, y, z) = g(x, y) - z$. Si g es diferenciable, también lo será f . Además

$$\frac{\partial f}{\partial x}(x, y, z) = \frac{\partial g}{\partial x}(x, y), \quad \frac{\partial f}{\partial y}(x, y, z) = \frac{\partial g}{\partial y}(x, y), \quad \frac{\partial f}{\partial z}(x, y, z) = -1.$$

Entonces el plano tangente a la superficie en el punto (x_0, y_0, z_0) tiene por ecuación

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

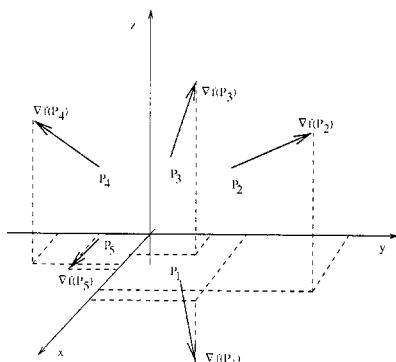
que podemos escribir como

$$(z - z_0) = (x - x_0) \frac{\partial g}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial g}{\partial y}(x_0, y_0).$$

Con frecuencia nos referiremos a $\vec{\nabla} f$ como **campo vectorial gradiente**, por el hecho de que $\vec{\nabla} f$ asigna un vector a cada punto del dominio de f . Su gráfica sería un subconjunto de \mathbb{R}^6 , pero se suele representar $\vec{\nabla} f(P)$, para cada punto P , como un vector que sale del punto P en lugar del origen.

El campo vectorial gradiente tiene un importante significado geométrico: muestra la dirección en la cual f crece más rápido y la dirección que es ortogonal a las superficies (o curvas, en el plano) de nivel de f .

10.5. GRADIENTES Y DERIVADAS DIRECCIONALES.



Ejemplo 1. Calcular la ecuación del plano tangente a la superficie definida por $3xy^2 + 2xz^2 = 5$ en $(1, 1, 1)$.

Solución:

Aquí $f(x, y, z) = 3xy^2 + 2xz^2$ y $\vec{\nabla} f = (3y^2 + 2z^2, 6xy, 4xz)$, que en el punto $(1, 1, 1)$ es el vector $(5, 6, 4)$. Así, el plano tangente es

$$(5, 6, 4) \cdot (x - 1, y - 1, z - 1) = 0 \quad \iff \quad 5x + 6y + 4z = 15.$$

Ejemplo 2. ¿En qué puntos de la superficie $z = 3xy - x^3 - y^3$ el plano tangente es horizontal?

Solución:

La función $g(x, y) = 3xy - x^3 - y^3$ tiene como derivadas parciales primeras

$$\frac{\partial g}{\partial x}(x, y) = 3y - 3x^2, \quad \frac{\partial g}{\partial y}(x, y) = 3x - 3y^2.$$

Estas se anulan a la vez sólo en $(0, 0)$ y $(1, 1)$. Luego la superficie tiene un plano tangente horizontal sólo en los puntos $(0, 0, 0)$ y $(1, 1, 1)$.

Ejemplo 3. Hallar un vector unitario normal a la superficie S dada por $z = x^2y^2 + xy + y + 2$ en el punto $(0, 0, 2)$.

Solución:

Sea $f(x, y, z) = x^2y^2 + xy + y + 2 - z$, y consideremos la superficie S definida por $f(x, y, z) = 0$. El gradiente está dado por

$$\vec{\nabla} f(x, y, z) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k} = (2xy^2 + y) \vec{i} + (2x^2y + x + 1) \vec{j} - \vec{k},$$

de manera que

$$\vec{\nabla} f(0, 0, 2) = \vec{j} - \vec{k}.$$

Este vector es perpendicular a S en el punto $(0, 0, 2)$ y, para hallar una normal unitaria \vec{n} , dividimos este vector entre su longitud para obtener

$$\vec{n} = \frac{\vec{\nabla} f(0, 0, 2)}{\|\vec{\nabla} f(0, 0, 2)\|} = \frac{1}{\sqrt{2}} [\vec{j} - \vec{k}].$$

Ejercicios.

1) Para las funciones $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ y $\vec{g}: \mathbb{R} \rightarrow \mathbb{R}^3$, halla $\vec{\nabla} f$ y \vec{g}' y evalúa $(f \circ \vec{g})'(1)$.

- a) $f(x, y, z) = xz + yz + xy$, $g(t) = (e^t, \cos t, \sin t)$;
- b) $f(x, y, z) = \exp(xyz)$, $g(t) = (6t, 3t^2, t^3)$;
- c) $f(x, y, z) = (x^2 + y^2 + z^2) \log(x^2 + y^2 + z^2)^{1/2}$, $g(t) = (e^t, e^{-t}, t)$.

2) Calcula la derivada direccional de f en el punto \vec{P} y en la dirección dados.

- a) $f(x, y) = \log \sqrt{x^2 + y^2}$, $\vec{P} = (1, 0)$, $\vec{v} = \frac{1}{\sqrt{5}} (2\vec{i} + \vec{j})$;
- b) $f(x, y) = xy^2 + x^3y$, $\vec{P} = (4, -2)$, $\vec{v} = \frac{1}{\sqrt{10}} (\vec{i} + 3\vec{j})$;
- c) $f(x, y) = e^x \cos(\pi y)$, $\vec{P} = (0, -1)$, $\vec{v} = -\frac{1}{\sqrt{5}} (\vec{i} - 2\vec{j})$;
- d) $f(x, y, z) = xy^2 + y^2z^3 + z^3x$, $\vec{P} = (4, -2, -1)$, $\vec{v} = \frac{1}{\sqrt{14}} (\vec{i} + 3\vec{j} + 2\vec{k})$;
- e) $f(x, y, z) = x^{yz}$, $\vec{P} = (e, e, 0)$, $\vec{v} = \frac{1}{13} (12\vec{i} + 3\vec{j} + 4\vec{k})$;
- f) $f(x, y, z) = xy + yz + zx$, $\vec{P} = (1, 1, 2)$, $\vec{v} = (10, -1, 2)$;

3) Si $h(x, y) = 2 \exp(-x^2) + \exp(-3y^2)$ denota la altura de una montaña en la posición (x, y) , desde el punto $(1, 1)$, ¿en qué dirección debería uno a comenzar a caminar para subir más rápido?

4) Calcula los gradientes de las funciones siguientes.

- a) $f(x, y, z) = x \exp(-x^2 - y^2 - z^2)$;
- b) $f(x, y, z) = z^2 e^x \cos y$;
- c) $f(x, y, z) = \frac{xyz}{x^2 + y^2 + z^2}$;
- d) $f(x, y, z) = x^2 + y^2 - z^2$;
- e) $f(x, y, z) = (x+z)e^{x-y}$;
- f) $f(x, y, z) = \log(x^2 + y^2 + z^2)$.

5) Calcula el plano tangente a $(1, 0, 1)$ para cada una de las funciones del ejercicio anterior.

6) Halla la ecuación del plano tangente a $z = x^2 + 2y^3$ en $(1, 1, 3)$.

10.6 Derivadas parciales iteradas.

El objetivo principal de esta sección es probar un teorema que asegura la igualdad de las segundas derivadas parciales mixtas de una función. Recordemos que si $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ es de clase \mathcal{C}^1 quiere decir que las derivadas parciales $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ y $\frac{\partial f}{\partial z}$ existen y son continuas, y por el teorema 10.5 la función f es diferenciable. Si estas derivadas, a su vez, tienen derivadas parciales continuas, decimos que f es de clase \mathcal{C}^2 , o que es dos veces continuamente diferenciable. De la misma manera, si decimos que f es de clase \mathcal{C}^3 , significa que f tiene derivadas parciales iteradas continuas de tercer orden, y así sucesivamente.

En el caso de una función de dos variables, $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, si las derivadas primeras $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ son continuamente diferenciables, al tomar las derivadas parciales segundas, obtenemos las cuatro funciones

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right), \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right), \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right).$$

Todas estas derivadas se llaman **derivadas parciales iteradas**, y las dos últimas se llaman **derivadas parciales mixtas**. De manera parecida, en el caso de una función de tres variables, $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, si las derivadas primeras $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ y $\frac{\partial f}{\partial z}$ son continuamente diferenciables, al tomar las derivadas parciales segundas, obtenemos las nueve funciones

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial x \partial z}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}, \quad \frac{\partial^2 f}{\partial y \partial z}, \quad \frac{\partial^2 f}{\partial z \partial x}, \quad \frac{\partial^2 f}{\partial z \partial y}, \quad \frac{\partial^2 f}{\partial z^2}.$$

Ejemplo 1. Halla las segundas derivadas parciales de $f(x, y) = \sin x^2 y$.

Solución:

Las derivadas primeras son

$$\frac{\partial f}{\partial x} = 2xy \cos x^2 y, \quad \frac{\partial f}{\partial y} = x^2 \cos x^2 y.$$

Las derivadas parciales de segundo orden son

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= 2y \cos x^2 y - 4x^2 y^2 \sin x^2 y, \\ \frac{\partial^2 f}{\partial y^2} &= -x^4 \sin x^2 y \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = 2x \cos x^2 y - 2x^3 y \sin x^2 y, \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = 2x \cos x^2 y - 2x^3 y \sin x^2 y. \end{aligned}$$

Ejemplo 2. Halla las segundas derivadas parciales de $f(x, y) = \ln(x^2 + y^3)$.

Solución:

Las derivadas primeras son

$$\frac{\partial f}{\partial x} = \frac{2x}{x^2 + y^3}, \quad \frac{\partial f}{\partial y} = \frac{3y^2}{x^2 + y^3}.$$

Las derivadas parciales de segundo orden son

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{2(x^2 + y^3) - 2x(2x)}{(x^2 + y^3)^2} = \frac{2(y^3 - x^2)}{(x^2 + y^3)^2}, \\ \frac{\partial^2 f}{\partial y^2} &= \frac{6y(x^2 + y^3) - 3y^2(3y^2)}{(x^2 + y^3)^2} = \frac{3y(2x^2 - y^3)}{(x^2 + y^3)^2}, \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{-3y^2(2x)}{(x^2 + y^3)^2} = -\frac{6xy^2}{(x^2 + y^3)^2}, \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{-2x(3y^2)}{(x^2 + y^3)^2} = -\frac{6xy^2}{(x^2 + y^3)^2}. \end{aligned}$$

En todos estos ejemplos nótese que los pares de derivadas parciales mixtas son iguales. Dado que en ninguno de los dos casos f era simétrica en x e y , esta igualdad de las derivadas parciales mixtas no puede deberse a la simetría. En realidad se debe a la continuidad, y es un hecho básico. Lo probaremos en el siguiente teorema para funciones de dos variables, pero la demostración se puede extender con facilidad a funciones de n variables.

Teorema 10.10 Si $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ es de clase C^2 , es decir dos veces continuamente diferenciable, entonces las derivadas parciales mixtas son iguales. Esto es

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

Demostración:

Consideremos la siguiente expresión

$$S(\Delta x, \Delta y) = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) - f(x_0, y_0 + \Delta y) + f(x_0, y_0).$$

Si mantenemos y_0 e Δy fijos, podemos definir

$$g(x) = f(x, y_0 + \Delta y) - f(x, y_0)$$

como función de una sola variable, de modo que resulta

$$S(\Delta x, \Delta y) = g(x_0 + \Delta x) - g(x_0).$$

10.6. DERIVADAS PARCIALES ITERADAS.

Por el teorema del valor medio para funciones de una variable, esta última expresión es igual a $g'(\bar{x}) \Delta x$, para algún \bar{x} entre x_0 y $x_0 + \Delta x$. Entonces,

$$S(\Delta x, \Delta y) = g(x_0 + \Delta x) - g(x_0) = g'(\bar{x}) \Delta x = \left[\frac{\partial f}{\partial x}(\bar{x}, y_0 + \Delta y) - \frac{\partial f}{\partial x}(\bar{x}, y_0) \right] \Delta x.$$

Aplicando nuevamente el teorema del valor medio, obtenemos

$$S(\Delta x, \Delta y) = \frac{\partial^2 f}{\partial y \partial x}(\bar{x}, \bar{y}) \Delta x \Delta y.$$

Como $\frac{\partial^2 f}{\partial y \partial x}$ es continua, tenemos que

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \lim_{(\Delta x, \Delta y) \rightarrow (0,0)} \frac{S(\Delta x, \Delta y)}{\Delta x \Delta y}.$$

De manera análoga, se muestra que $\frac{\partial^2 f}{\partial x \partial y}$ está dada por la misma fórmula de límite, lo cual prueba el resultado.

c.q.d.

Si $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ es de clase C^3 , es decir tres veces continuamente diferenciable en las tres variables x , y y z , entonces podemos calcular las derivadas parciales iteradas en el orden que nos plazca dado que las derivadas parciales mixtas son iguales:

$$\frac{\partial^3 f}{\partial x \partial y \partial z} = \frac{\partial^3 f}{\partial x \partial z \partial y} = \frac{\partial^3 f}{\partial z \partial y \partial x} = \frac{\partial^3 f}{\partial z \partial x \partial y} = \frac{\partial^3 f}{\partial y \partial x \partial z} = \frac{\partial^3 f}{\partial y \partial z \partial x}.$$

A veces se usa la notación f_x , f_y , f_z para las derivadas parciales primeras: $f_x = \frac{\partial f}{\partial x}$, etc.

Con esta notación, escribimos $f_{xy} = (f_x)_y = \frac{\partial^2 f}{\partial y \partial x}$, de manera que se invierte el orden de x e y en las dos notaciones; afortunadamente, la igualdad de las derivadas parciales mixtas hace irrelevante esta potencial ambigüedad.

Ejemplo 1. Verifica la igualdad de las segundas derivadas parciales mixtas para la función

$$f(x, y, z) = x e^y \operatorname{sen} \pi z.$$

Solución:

Las derivadas primeras son

$$\frac{\partial f}{\partial x} = e^y \operatorname{sen} \pi z, \quad \frac{\partial f}{\partial y} = x e^y \operatorname{sen} \pi z, \quad \frac{\partial f}{\partial z} = \pi x e^y \cos \pi z.$$

Entonces para las derivadas segundas obtenemos

$$\frac{\partial^2 f}{\partial x^2} = 0, \quad \frac{\partial^2 f}{\partial y^2} = x e^y \operatorname{sen} \pi z, \quad \frac{\partial^2 f}{\partial z^2} = -\pi^2 x e^y \operatorname{sen} \pi z,$$

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x} &= e^y \operatorname{sen} \pi z = \frac{\partial^2 f}{\partial x \partial y}, \\ \frac{\partial^2 f}{\partial z \partial x} &= \pi e^y \cos \pi z = \frac{\partial^2 f}{\partial x \partial z}, \\ \frac{\partial^2 f}{\partial y \partial z} &= \pi x e^y \cos \pi z = \frac{\partial^2 f}{\partial z \partial y}.\end{aligned}$$

Ejemplo 2. Sea $z = f(x, y)$ con $x = g(s, t)$ e $y = h(s, t)$. Luego se puede pensar en una nueva función $k(s, t) = f(g(s, t), h(s, t))$. Calcula k_{st} usando la regla de la cadena y comprueba que $k_{st} = k_{ts}$.

Solución:

Por la regla de la cadena

$$k_s = f_x g_s + f_y h_s \quad \text{y} \quad k_{st} = (k_s)_t = (f_x)_t g_s + f_x (g_s)_t + (f_y)_t h_s + f_y (h_s)_t.$$

Aplicando la regla de la cadena a $(f_x)_t$ y a $(f_y)_t$ se tiene

$$(f_x)_t = f_{xx} g_t + f_{xy} h_t \quad \text{y} \quad (f_y)_t = f_{yx} g_t + f_{yy} h_t$$

y sustituyendo resulta

$$\begin{aligned}k_{st} &= (f_{xx} g_t + f_{xy} h_t) g_s + f_x g_{st} + (f_{yx} g_t + f_{yy} h_t) h_s + f_y h_{st} \\ &= f_{xx} g_t g_s + f_{xy} h_t g_s + f_x g_{st} + f_{yx} g_t h_s + f_{yy} h_t h_s + f_y h_{st} \\ &= f_{xx} g_t g_s + f_{xy} (h_t g_s + g_t h_s) + f_{yy} h_t h_s + f_x g_{st} + f_y h_{st}.\end{aligned}$$

De la misma manera obtenemos

$$k_t = f_x g_t + f_y h_t \quad \text{y} \quad k_{ts} = (k_t)_s = (f_x)_s g_t + f_x (g_t)_s + (f_y)_s h_s + f_y (h_t)_s.$$

Aplicando la regla de la cadena a $(f_x)_s$ y a $(f_y)_s$ se tiene

$$(f_x)_s = f_{xx} g_s + f_{xy} h_s \quad \text{y} \quad (f_y)_s = f_{yx} g_s + f_{yy} h_s$$

y sustituyendo resulta

$$\begin{aligned}k_{ts} &= (f_{xx} g_s + f_{xy} h_s) g_t + f_x g_{ts} + (f_{yx} g_s + f_{yy} h_s) h_t + f_y h_{ts} \\ &= f_{xx} g_s g_t + f_{xy} h_s g_t + f_x g_{ts} + f_{yx} g_s h_t + f_{yy} h_s h_t + f_y h_{ts} \\ &= f_{xx} g_s g_t + f_{xy} (h_s g_t + g_s h_t) + f_{yy} h_s h_t + f_x g_{ts} + f_y h_{ts}.\end{aligned}$$

Claramente las dos expresiones coinciden si las funciones $g(s, t) = x$ y $h(s, t) = y$ son funciones de clase \mathcal{C}^2 (dos veces continuamente diferenciables).

Ejercicios.

- 1) Calcula las segundas derivadas parciales para cada una de las siguientes funciones.

10.6. DERIVADAS PARCIALES ITERADAS.

- a) $f(x, y, z) = \exp(z) + \frac{1}{x} + x \exp(-y)$; b) $f(x, y) = \cos(xy^2)$;
 c) $f(x, y) = \log(x^2 + y)$; d) $f(x, y) = x \arctan \frac{x}{y}$.

2) Muestra que $u(x, y) = \exp(x)$ sen y satisface la ecuación (ecuación de Laplace)
 $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$.

3) Prueba que si $f_{xx} = \frac{\partial^2 f}{\partial x^2}$ etc..., las funciones

- a) $f(x, y) = \log(x^2 + y^2)$,
 b) $g(x, y, z) = \frac{1}{(x^2 + y^2 + z^2)^{\frac{1}{2}}}$,
 c) $h(x, y, z, w) = \frac{1}{x^2 + y^2 + z^2 + w^2}$,

satisfacen las respectivas ecuaciones de Laplace:

- a) $f_{xx} + f_{yy} = 0$,
 b) $g_{xx} + g_{yy} + g_{zz} = 0$,
 c) $h_{xx} + h_{yy} + h_{zz} + h_{ww} = 0$.

4) Sean $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ y $\vec{g}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dos campos vectoriales definidos del siguiente modo

$$\vec{f}(x, y) = e^{x-2y} \vec{i} + \text{sen}(2x + y) \vec{j},$$

$$\vec{g}(u, v, w) = (u + 2v^2 + 3w^3) \vec{i} + (2v - u^2) \vec{j},$$

- a) calcula cada una de las matrices jacobianas $D\vec{f}(x, y)$ y $D\vec{g}(u, v, w)$;
 b) calcula la función compuesta $\vec{h}(u, v, w) = \vec{f}(\vec{g}(u, v, w))$;
 c) calcula la matriz jacobiana $D\vec{h}(1, -1, 1)$.

5) Sean $\vec{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ y $\vec{g}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dos campos vectoriales definidos del siguiente modo

$$\vec{f}(x, y, z) = (x^2 + y + z) \vec{i} + (2x + y + z^2) \vec{j},$$

$$\vec{g}(u, v, w) = uv^2 w^2 \vec{i} + w^2 \text{sen } v \vec{j} + u^2 e^v \vec{k},$$

- a) calcula cada una de las matrices jacobianas $D\vec{f}(x, y, z)$ y $D\vec{g}(u, v, w)$;
 b) calcula la función compuesta $\vec{h}(u, v, w) = \vec{f}(\vec{g}(u, v, w))$;
 c) calcula la matriz jacobiana $D\vec{h}(u, 0, w)$.

6) Calcula todas las derivadas parciales mixtas de orden dos y orden tres para las siguientes funciones.

- | | |
|--|---|
| a) $f(x, y) = \frac{2x^2 + 7x^2y}{3xy}, \quad (x, y) \neq (0, 0);$ | b) $f(x, y) = 3x^2 + 2y^2;$ |
| c) $f(x, y) = \text{sen}(x^2 - 3xy);$ | d) $f(x, y) = x^2y^2e^{2xy};$ |
| e) $f(x, y, z) = x^2y + xy^2 + yz^2;$ | f) $f(x, y, z) = x^4y^3 - x^8 + y^4;$ |
| g) $f(x, y, z) = ze^{xy} + yz^3x^2;$ | h) $f(x, y, z) = e^{xyz} \text{sen}(xz).$ |

7) Verifica que $f_{xzw} = f_{zwx}$ para $f(x, y, z, w) = e^{xyz} \text{sen}(xw)$.

8) Comprueba que la función $g(x, t) = 2 + e^{-t} \text{sen } x$ (que representa la temperatura de una varilla de metal en la posición x y tiempo t) satisface la ecuación del calor $g_t = g_{xx}$. Dibuja la gráfica de g para tiempos positivos ($t \geq 0$). ¿Qué ocurre en el límite $t \rightarrow \infty$?

10.7 Trayectorias y velocidad. Longitud de arco.

Con frecuencia consideramos una curva como una línea trazada sobre un papel, tal como una línea recta o un círculo. Para tratar de manera efectiva con estos objetos, resulta conveniente pensar en una curva en \mathbb{R}^n , como el conjunto de valores de una función que manda un intervalo en \mathbb{R} a \mathbb{R}^n . A dicha función le llamaremos *trayectoria*. Una trayectoria está en el plano si $n = 2$, y en el espacio si $n = 3$.

Definición. Una *trayectoria* en \mathbb{R}^n es una función $\vec{\sigma}: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^n$. Si $\vec{\sigma}$ es diferenciable, decimos que $\vec{\sigma}$ es una *trayectoria diferenciable*. Si $\vec{\sigma}$ es de clase C^1 , decimos que $\vec{\sigma}$ es una C^1 *trayectoria*. Los puntos $\vec{\sigma}(a)$ y $\vec{\sigma}(b)$ se llaman **extremos** de la trayectoria. La imagen de $\vec{\sigma}$ se llama **curva** de $\vec{\sigma}$.

Es útil denotar la variable como t y pensar que $\vec{\sigma}(t)$ va trazando una curva en \mathbb{R}^n conforme t va variando. Si $\vec{\sigma}(t)$ es una trayectoria de \mathbb{R}^3 , podemos escribir $\vec{\sigma}(t) = (x(t), y(t), z(t))$ y llamamos a $x(t), y(t), z(t)$, **funciones componentes** de $\vec{\sigma}$. De manera similar podemos formar funciones componentes en \mathbb{R}^2 y en general en \mathbb{R}^n .

Ejemplo 1. Sean $f_1(t) = x_0 + v_1 t$, $f_2(t) = y_0 + v_2 t$ y $f_3(t) = z_0 + v_3 t$, tres funciones escalares. Podemos entonces formar la función vectorial

$$\vec{f}(t) = f_1(t)\vec{i} + f_2(t)\vec{j} + f_3(t)\vec{k} = (x_0 + v_1 t)\vec{i} + (y_0 + v_2 t)\vec{j} + (z_0 + v_3 t)\vec{k}.$$

Si v_1, v_2 y v_3 no son todos nulos, el radio vector $\vec{f}(t)$ recorre la recta que pasa por el punto $P(x_0, y_0, z_0)$ y tiene como componentes de dirección v_1, v_2 y v_3 , como vimos en la sección 9.1.

Ejemplo 2. A partir de las funciones $f_1(t) = \cos t$, $f_2(t) = \text{sen } t$ y $f_3(t) = 0$ podemos formar la función vectorial

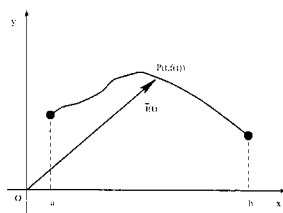
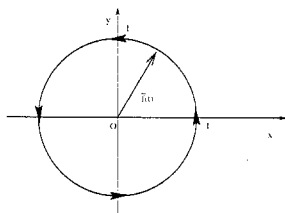
$$\vec{f}(t) = f_1(t)\vec{i} + f_2(t)\vec{j} + f_3(t)\vec{k} = \cos t\vec{i} + \text{sen } t\vec{j}.$$

10.7. TRAYECTORIAS Y VELOCIDAD. LONGITUD DE ARCO.

Para todo t resulta

$$\|\vec{f}(t)\| = \sqrt{\cos^2 t + \sin^2 t} = 1.$$

Dado que la tercera componente es cero, el radio vector $\vec{f}(t)$ está en el plano xy . Cuando t crece, el extremo de $\vec{f}(t)$ recorre el círculo unidad en el sentido contrario al de las agujas del reloj, efectuando una revolución completa cuando t aumenta en 2π .



Ejemplo 3. A toda función real definida en un cierto intervalo $[a, b]$ se le puede asociar de manera natural una función vectorial \vec{f} . Haciendo $f_1(t) = t$, $f_2(t) = f(t)$, $f_3(t) = 0$ obtenemos la función vectorial

$$\vec{f}(t) = t\vec{i} + f(t)\vec{j}.$$

Cuando t varía desde a hasta b , el radio vector $\vec{f}(t)$ recorre la gráfica de f de izquierda a derecha.

Por lo general, los objetos que se mueven en el espacio lo hacen en curvas suaves, por ejemplo no desaparecen, ni cambian repentinamente de velocidad. Así, restringiremos nuestro estudio a trayectorias suficientemente suaves, digamos \mathcal{C}^1 .

En la sección 10.3 aprendimos que si $\vec{\sigma}$ es una trayectoria en \mathbb{R}^3 , es decir, una función vectorial $\vec{\sigma}: \mathbb{R} \rightarrow \mathbb{R}^3$, su derivada $D\vec{\sigma}(t)$ es una matriz de 3×1 :

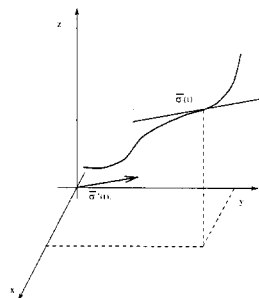
$$D\vec{\sigma}(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}.$$

Sea $\vec{\sigma}'(t)$ el correspondiente vector (fila):

$$\vec{\sigma}'(t) = (x'(t), y'(t), z'(t))$$

de modo que

$$\vec{\sigma}'(t) = \lim_{h \rightarrow 0} \frac{\vec{\sigma}(t_0 + h) - \vec{\sigma}(t_0)}{h}.$$



Intuitivamente podemos argumentar que el vector $\vec{\sigma}'(t_0)$ deberá ser paralelo a la recta L , tangente a la trayectoria $\vec{\sigma}$ en el punto $\vec{\sigma}(t_0)$, y que deberá representar la velocidad del objeto. En efecto, el cociente

$$\frac{\vec{\sigma}(t_0 + h) - \vec{\sigma}(t_0)}{h} = \frac{\vec{\sigma}(t_0 + h) - \vec{\sigma}(t_0)}{(t_0 + h) - t_0}$$

representa la velocidad dirigida promedio en el intervalo de tiempo de t_0 a $t_0 + h$ (desplazamiento total entre el tiempo transcurrido). Por lo tanto, cuando $h \rightarrow 0$, esta expresión tiende al vector de velocidad instantánea. Esto conduce a la siguiente definición.

Definición. Sea $\vec{\sigma}: \mathbb{R} \rightarrow \mathbb{R}^3$ una trayectoria de clase C^1 . El **vector velocidad** en $\vec{\sigma}(t)$ está dado por $\vec{v}(t) = \vec{\sigma}'(t) = (x'(t), y'(t), z'(t))$, y la **rapidez** del objeto está dada por $S(t) = \|\vec{\sigma}'(t)\|$, la longitud del vector $\vec{\sigma}'(t)$.

Como el vector velocidad $\vec{\sigma}'(t_0)$ es paralelo a la recta L tangente a la trayectoria $t \rightsquigarrow \vec{\sigma}(t)$ en el punto $\vec{\sigma}(t_0)$, una ecuación de la recta L tangente a $\vec{\sigma}(t)$ en $\vec{\sigma}(t_0)$ deberá venir dada en forma paramétrica por la fórmula

$$\vec{l}(\lambda) = \vec{\sigma}(t_0) + \lambda \vec{\sigma}'(t_0)$$

donde el parámetro λ varía entre los números reales.

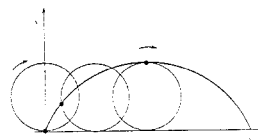
Ejemplo 1. Si $\vec{\sigma}: t \rightsquigarrow (\cos t, \sin t, t)$, entonces el vector velocidad es $\vec{v}(t) = \vec{\sigma}'(t) = (-\sin t, \cos t, 1)$. La rapidez de un punto es la magnitud de la velocidad

$$S(t) = \|\vec{v}(t)\| = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{2}.$$

Así, el punto se mueve con rapidez constante, aunque su velocidad no sea constante, pues cambia continuamente de dirección. La trayectoria del punto cuyo movimiento está dado por esta función $\vec{\sigma}$ se llama hélice (circular recta), que está sobre un cilindro circular recto.

Ejemplo 2. La función $\vec{\sigma}: t \rightsquigarrow (t - \sin t, 1 - \cos t)$ describe la función de posición de un punto en un círculo de radio 1, que va rodando.

El círculo está en el plano xy y rueda a lo largo del eje x con rapidez constante; esto es, el centro del círculo se mueve hacia la derecha a lo largo de la recta $y = 1$, con rapidez constante de 1 radián por unidad de tiempo. El movimiento del punto $\vec{\sigma}(t)$ es más complicado; su curva se conoce como cicloide.



Definición. Dada una partícula moviéndose sobre una trayectoria $\vec{\sigma}(t)$, se define la tasa de cambio del vector velocidad como **aceleración**. Así,

$$\vec{a}(t) = \vec{\sigma}''(t) = (x''(t), y''(t), z''(t)).$$

Si un objeto de masa m se mueve en \mathbb{R}^3 , la fuerza \vec{F} que actúa sobre ella en el punto $\vec{\sigma}(t)$, se relaciona con la aceleración por medio de la segunda ley de Newton:

$$\vec{F}(\vec{\sigma}(t)) = m \cdot \vec{a}(t).$$

En particular, si no actúa fuerza alguna sobre una partícula, $\vec{a}(t) = \vec{0}$, de modo que $\vec{\sigma}'(t)$ es constante y la partícula sigue una recta.

10.7. TRAYECTORIAS Y VELOCIDAD. LONGITUD DE ARCO.

Definición. Sea $\vec{\sigma}: [a, b] \rightarrow \mathbb{R}^n$ una trayectoria de clase C^1 . La **longitud** de $\vec{\sigma}$ está definida como

$$l(\vec{\sigma}) = \int_a^b \|\vec{\sigma}'(t)\| dt. \quad (10.3)$$

En \mathbb{R}^2 y \mathbb{R}^3 la fórmula se reduce, respectivamente, a

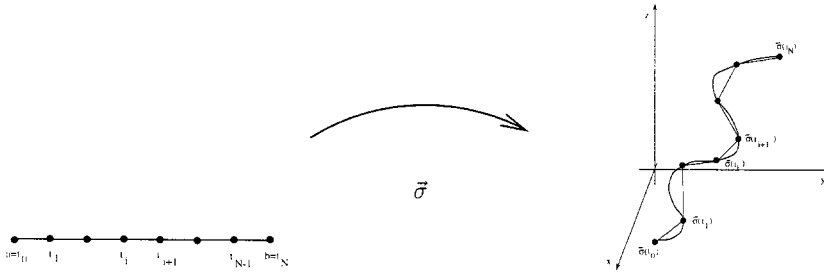
$$l(\vec{\sigma}) = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2} dt,$$

$$l(\vec{\sigma}) = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2} dt.$$

Para justificar esta definición en \mathbb{R}^3 , podemos proceder de la siguiente manera. Partimos el intervalo $[a, b]$ en N partes iguales,

$$a = t_0 < t_1 < \dots < t_N = b, \quad \text{con} \quad \Delta t = t_{i+1} - t_i = \frac{b-a}{N},$$

para $0 \leq i \leq N-1$. Después consideramos la poligonal obtenida al unir pares sucesivos de puntos $\vec{\sigma}(t_i)$, $\vec{\sigma}(t_{i+1})$, para $0 \leq i \leq N-1$. Esto produce una aproximación poligonal a $\vec{\sigma}$ como muestra la figura.



De la fórmula de la distancia en \mathbb{R}^3 se sigue que el segmento que une $\vec{\sigma}(t_i)$ a $\vec{\sigma}(t_{i+1})$ tiene una longitud

$$\|\vec{\sigma}(t_{i+1}) - \vec{\sigma}(t_i)\| = \sqrt{[x(t_{i+1}) - x(t_i)]^2 + [y(t_{i+1}) - y(t_i)]^2 + [z(t_{i+1}) - z(t_i)]^2}.$$

Aplicando el teorema del valor medio a las funciones $x(t)$, $y(t)$, $z(t)$ en $[t_i, t_{i+1}]$, obtenemos tres puntos t_i^* , t_i^{**} , t_i^{***} , tales que

$$x(t_{i+1}) - x(t_i) = x'(t_i^*) (t_{i+1} - t_i),$$

$$y(t_{i+1}) - y(t_i) = y'(t_i^{**}) (t_{i+1} - t_i)$$

y

$$z(t_{i+1}) - z(t_i) = z'(t_i^{***}) (t_{i+1} - t_i).$$

Luego el segmento que une $\vec{\sigma}(t_i)$ a $\vec{\sigma}(t_{i+1})$ tiene una longitud

$$\|\vec{\sigma}(t_{i+1}) - \vec{\sigma}(t_i)\| = \sqrt{[x'(t_i^*)]^2 + [y'(t_i^{**})]^2 + [z'(t_i^{***})]^2} (t_{i+1} - t_i).$$

Entonces la longitud de nuestra poligonal de aproximación es

$$S_N = \sum_{i=0}^{N-1} \sqrt{[x'(t_i^*)]^2 + [y'(t_i^{**})]^2 + [z'(t_i^{***})]^2} (t_{i+1} - t_i).$$

Cuando $N \rightarrow \infty$, esta poligonal aproxima mejor la imagen de la curva $\vec{\sigma}$. Por lo tanto definimos la longitud de arco como el límite, si es que existe, de la sucesión S_N cuando $N \rightarrow \infty$. Como se supone que las derivadas x' , y' , z' son continuas en $[a, b]$, podemos concluir que el límite existe y está dado por

$$l(\vec{\sigma}) = \lim_{N \rightarrow \infty} S_N = \int_a^b \sqrt{[x'(t)]^2 + [y'(t)]^2 + [z'(t)]^2} dt.$$

De manera parecida se procede en el caso de \mathbb{R}^2 , y se puede extender al caso de \mathbb{R}^n para llegar a la fórmula (10.3)

$$l(\vec{\sigma}) = \int_a^b \|\vec{\sigma}'(t)\| dt.$$

Ejemplo 1. La longitud de arco de la curva $\vec{\sigma}(t) = (r \cos t, r \sin t)$, para $0 \leq t \leq 2\pi$, es

$$l(\vec{\sigma}) = \int_0^{2\pi} \sqrt{[x'(t)]^2 + [y'(t)]^2} dt = \int_0^{2\pi} \sqrt{[-r \sin t]^2 + [r \cos t]^2} dt = 2\pi r.$$

Si permitimos que $0 \leq t \leq 4\pi$, hubiéramos obtenido $l(\vec{\sigma}) = 4\pi r$, pues la trayectoria recorrería dos veces el mismo círculo.

Ejemplo 2. Halla la velocidad, rapidez y longitud de arco, del punto que traza la cicloide dada por $\vec{\sigma}: t \rightsquigarrow (t - \sin t, 1 - \cos t)$.

Solución:

El vector velocidad es

$$\vec{\sigma}'(t) = (1 - \cos t, \sin t)$$

y la rapidez del punto es

$$\|\vec{\sigma}'(t)\| = \sqrt{(1 - \cos t)^2 + \sin^2 t} = \sqrt{2 - 2 \cos t}.$$

Por lo tanto $\vec{\sigma}(t)$ se mueve con rapidez variable, aunque el círculo rueda con rapidez constante. Más aún, la rapidez de $\vec{\sigma}(t)$ es cero cuando t es un múltiplo entero de 2π . En estos valores, la coordenada y de $\vec{\sigma}(t)$ es cero, por lo que el punto está en el eje x .

La longitud de arco de un ciclo es

$$l(\vec{\sigma}) = \int_0^{2\pi} \sqrt{2 - 2 \cos t} dt = 2 \int_0^{2\pi} \sqrt{\frac{1 - \cos t}{2}} dt = 2 \int_0^{2\pi} \sin \frac{t}{2} dt = 8.$$

10.7. TRAYECTORIAS Y VELOCIDAD. LONGITUD DE ARCO.

Ejercicios.

- 1) Halla $\vec{\sigma}'(t)$ y $\vec{\sigma}'(0)$ en cada uno de los casos siguientes.
 - a) $\vec{\sigma}(t) = (\exp(t), \cos t, \text{sen } t)$;
 - b) $\vec{\sigma}(t) = (\text{sen}(2t), \log(1+t), t)$;
 - c) $\vec{\sigma}(t) = (\cos \alpha \cos(wt), \text{sen } \alpha \cos(wt), \text{sen}(wt))$.
- 2) Para cada una de las siguientes curvas, determina los vectores velocidad, aceleración y la ecuación de la recta tangente en el valor de t especificado.
 - a) $\vec{r}(t) = \sqrt{2}t\vec{i} + \exp(t)\vec{j} + \exp(-t)\vec{k}$, $t = 0$;
 - b) $\vec{r}(t) = t\vec{i} + t\vec{j} + \frac{2}{3}t^{3/2}\vec{k}$, $t = 9$;
 - c) $\vec{r}(t) = 2\cos(t)\vec{i} - 2\text{sen}(t)\vec{j} + 3t\vec{k}$, $t = \pi/2$;
 - d) $\vec{r}(t) = (\cos \alpha \cos(wt), \text{sen } \alpha \cos(wt), \text{sen}(wt))$, $t = \pi/4$;
 - e) $\vec{r}(t) = \cos(\pi t)\vec{i} + \text{sen}(\pi t)\vec{j} + t\vec{k}$, $t = 2$;
 - f) $\vec{r}(t) = a\vec{i} + t\vec{b} + t^2\vec{c}$, $t = -1$.
- 3) Demuestra que $\vec{r}(t) = at\vec{i} + bt^2\vec{j}$ parametriza una parábola. Halla una ecuación en x e y para dicha parábola.
- 4) Halla
 - a) los puntos de la curva $\vec{r}(t) = t\vec{i} + (1+t^2)\vec{j}$ en los cuales $\vec{r}(t)$ y $\vec{r}'(t)$ son perpendiculares;
 - b) los puntos en los cuales tienen el mismo sentido;
 - c) los puntos en los cuales tienen sentido opuesto.
- 5) Halla la longitud de las curvas
 - a) $\vec{r}(t) = a\cos t\vec{i} + a\text{sen } t\vec{j} + bt\vec{k}$, desde $t = 0$ hasta $t = 2\pi$;
 - b) $\vec{r}(t) = t\vec{i} + \frac{2}{3}\sqrt{2}t^{3/2}\vec{j} + \frac{2}{3}t^2\vec{k}$, desde $t = 0$ hasta $t = 2$;
 - c) $\vec{r}(t) = t^3\vec{i} + t^2\vec{j}$, desde $t = 0$ hasta $t = 1$.
- 6) Halla la longitud de arco de las trayectorias siguientes entre los puntos indicados.
 - a) $\vec{\sigma}(t) = (2t, t^2, \log t)$, definida para $t > 0$, entre $(2, 1, 0)$ y $(4, 4, \log 2)$;
 - b) $\vec{\sigma}(t) = (t, t \text{ sen } t, t \cos t)$, entre $(0, 0, 0)$ y $(\pi, 0, -\pi)$.

Capítulo 11

Derivadas de orden superior: máximos y mínimos.

En el cálculo de una variable, para saber si una función $f(x)$ tiene algún máximo o mínimo local, se buscan los puntos críticos, es decir los puntos para los cuales se anula la derivada primera, y en cada uno de dichos puntos se observa el signo de la segunda derivada. Si esta última es negativa entonces hay un máximo local, si es positiva hay un mínimo local y si es nula el criterio falla. En este capítulo se extienden estos métodos a funciones de varias variables con valores reales.

11.1 Teorema de Taylor.

Para funciones suaves de una variable $f: \mathbb{R} \rightarrow \mathbb{R}$, el teorema de Taylor (ver sección 2.3) asegura que

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots + \frac{f^{(k)}(a)}{k!}(x-a)^k + R_k(x, a)$$

donde el resto o residuo viene dado por

$$R_k(x, a) = \int_a^x \frac{(x-t)^k}{k!} f^{(k+1)}(t) dt.$$

Para x cerca de a , este error $R_k(x, a)$ es pequeño y de orden k , es decir $\lim_{x \rightarrow a} \frac{R_k(x, a)}{(x-a)^k} = 0$.

Queremos ahora probar un teorema análogo válido para funciones de varias variables. Ya conocemos una versión del primer orden si consideramos la definición de diferenciabilidad dada en la sección 10.3: si $f: \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable en \vec{x}_0 entonces

$$\lim_{\vec{x} \rightarrow \vec{x}_0} \frac{|R_1(\vec{x}, \vec{x}_0)|}{\|\vec{x} - \vec{x}_0\|} = 0, \quad \text{donde } R_1(\vec{x}, \vec{x}_0) = f(\vec{x}) - f(\vec{x}_0) - [\mathbf{D}f(\vec{x}_0)](\vec{x} - \vec{x}_0),$$

de modo que

$$f(\vec{x}) = f(\vec{x}_0) + [\mathbf{D}f(\vec{x}_0)](\vec{x} - \vec{x}_0) + R_1(\vec{x}, \vec{x}_0).$$

Si denotamos con $\vec{h} = \vec{x} - \vec{x}_0$ y $R_1(\vec{x}, \vec{x}_0) = R_1(\vec{h}, \vec{x}_0)$ tendremos

Teorema 11.1 Fórmulas de Taylor de primer y segundo orden.
 Sea $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciable en $\vec{x}_0 \in \mathbb{R}^n$. Entonces podemos escribir

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\vec{x}_0) + R_1(\vec{h}, \vec{x}_0), \quad (11.1)$$

$$\text{con } \lim_{\vec{h} \rightarrow \vec{0}} \frac{|R_1(\vec{h}, \vec{x}_0)|}{\|\vec{h}\|} = 0.$$

Si las derivadas parciales son continuas hasta el tercer orden, obtenemos

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\vec{x}_0) + \frac{1}{2} \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) + R_2(\vec{h}, \vec{x}_0), \quad (11.2)$$

$$\text{con } \lim_{\vec{h} \rightarrow \vec{0}} \frac{|R_2(\vec{h}, \vec{x}_0)|}{\|\vec{h}\|^2} = 0.$$

Demostración:

Por la regla de la cadena tenemos

$$\frac{d}{dt} f(\vec{x}_0 + t\vec{h}) = \left[\mathbf{D}f(\vec{x}_0 + t\vec{h}) \right] \vec{h} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{x}_0) h_i.$$

Ahora integramos ambos lados de la igualdad de $t = 0$ a $t = 1$:

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{x}_0 + t\vec{h}) h_i dt = \sum_{i=1}^n \int_0^1 \frac{\partial f}{\partial x_i}(\vec{x}_0 + t\vec{h}) h_i dt.$$

Si ahora integramos por partes, con la fórmula general $\int_0^1 u \frac{dv}{dt} dt = - \int_0^1 v \frac{du}{dt} dt + u v \Big|_0^1$,

en este caso podemos considerar las funciones $v = t - 1$ y $u = \frac{\partial f}{\partial x_i}(\vec{x}_0 + t\vec{h}) h_i$, luego

por un lado $u v \Big|_0^1 = (t - 1) \frac{\partial f}{\partial x_i}(\vec{x}_0 + t\vec{h}) h_i \Big|_0^1 = \frac{\partial f}{\partial x_i}(\vec{x}_0) h_i$ y por otro, por la regla de

la cadena, $\frac{du}{dt} = \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0 + t\vec{h}) h_i h_j$. Entonces, resulta

$$\sum_{i=1}^n \int_0^1 \frac{\partial f}{\partial x_i}(\vec{x}_0 + t\vec{h}) h_i dt = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{x}_0) h_i + \sum_{i,j=1}^n \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0 + t\vec{h}) h_i h_j dt.$$

11.1. TEOREMA DE TAYLOR.

Así, hemos probado la identidad (11.1)

$$f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{x}_0) h_i + R_1(\vec{h}, \vec{x}_0)$$

donde el residuo viene dado por la fórmula explícita

$$R_1(\vec{h}, \vec{x}_0) = \sum_{i,j=1}^n \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0 + t\vec{h}) h_i h_j dt. \quad (11.3)$$

El integrando es una función continua de t y por lo tanto está acotada en un entorno de \vec{x}_0 . Así, para una constante $M \geq 0$ obtenemos para $\|\vec{h}\|$ pequeña $|R_1(\vec{h}, \vec{x}_0)| \leq \|\vec{h}\|^2 M$ y entonces

$$\frac{|R_1(\vec{h}, \vec{x}_0)|}{\|\vec{h}\|} \leq \|\vec{h}\| M \rightarrow 0 \quad \text{cuando } \vec{h} \rightarrow \vec{0}.$$

Si ahora integramos por partes la expresión (11.3) para $R_1(\vec{h}, \vec{x}_0)$, considerando las funciones $v = -\frac{(t-1)^2}{2}$ y $u = \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0 + t\vec{h}) h_i h_j$ obtenemos

$$R_1(\vec{h}, \vec{x}_0) = \sum_{i,j,k=1}^n \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\vec{x}_0 + t\vec{h}) h_i h_j h_k dt + \\ + \sum_{i,j=1}^n \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0 + t\vec{h}) h_i h_j$$

de manera que hemos probado la identidad (11.2)

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\vec{x}_0) + \frac{1}{2} \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) + R_2(\vec{h}, \vec{x}_0)$$

donde el residuo viene dado por la fórmula explícita

$$R_2(\vec{h}, \vec{x}_0) = \sum_{i,j,k=1}^n \int_0^1 \frac{(t-1)^2}{2} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\vec{x}_0 + t\vec{h}) h_i h_j h_k dt. \quad (11.4)$$

Por el mismo razonamiento hecho para $R_1(\vec{h}, \vec{x}_0)$, obtenemos $|R_2(\vec{h}, \vec{x}_0)| \leq \|\vec{h}\|^3 M$, para $\|\vec{h}\|$ pequeño, entonces

$$\frac{|R_2(\vec{h}, \vec{x}_0)|}{\|\vec{h}\|^2} \leq \|\vec{h}\| M \rightarrow 0 \quad \text{cuando } \vec{h} \rightarrow \vec{0}.$$

c.q.d.

CAPÍTULO 11. DERIVADAS DE ORDEN SUPERIOR: MÁXIMOS Y MÍNIMOS.

Ejemplo. Calcula la fórmula de Taylor de segundo orden para $f(x, y) = \sin(x + 2y)$, alrededor del punto $\vec{x}_0 = (0, 0)$.

Solución:

Empezamos hallando

$$f(0, 0) = 0, \quad \frac{\partial f}{\partial x}(0, 0) = \cos(0 + 2 \cdot 0) = 1, \quad \frac{\partial f}{\partial y}(0, 0) = 2 \cos(0 + 2 \cdot 0) = 2,$$

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

Luego obtenemos

$$f(\vec{h}) = f(h_1, h_2) = h_1 + 2h_2 + R_2(\vec{h}, \vec{0})$$

donde $\frac{|R_2(\vec{h}, \vec{x}_0)|}{\|\vec{h}\|^2} \rightarrow 0$ cuando $\vec{h} \rightarrow \vec{0}$.

Ejercicio.

Determina la fórmula de Taylor de segundo orden para las funciones dadas respecto del punto dado.

- $f(x, y) = \frac{1}{(x^2 + y^2 + 1)}, \quad x_0 = 0, y_0 = 0;$
- $f(x, y) = \exp(-x^2 - y^2) \cos(xy), \quad x_0 = 0, y_0 = 0;$
- $f(x, y) = \sin(xy) + \cos(xy), \quad x_0 = 0, y_0 = 0;$
- $f(x, y) = \exp(x^2 - 2x + 1) \cos y, \quad x_0 = 0, y_0 = 0;$
- $f(x, y) = y^x, \quad x_0 = 1, y_0 = 1.$

11.2 Extremos de funciones con valores reales.

Definición. Si $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es una función escalar dada, un punto $\vec{x}_0 \in U$ se llama **mínimo local** de f si existe un entorno $E(\vec{x}_0, r)$ tal que para todos los puntos $\vec{x} \in E(\vec{x}_0, r)$, $f(\vec{x}) \geq f(\vec{x}_0)$. Se dirá **mínimo estricto** si $f(\vec{x}) > f(\vec{x}_0)$. De manera análoga, un punto $\vec{x}_0 \in U$ se llama **máximo local** de f si existe un entorno $E(\vec{x}_0, r)$ tal que para todos los puntos $\vec{x} \in E(\vec{x}_0, r)$, $f(\vec{x}) \leq f(\vec{x}_0)$. Se dirá **máximo estricto** si $f(\vec{x}) < f(\vec{x}_0)$. El punto $\vec{x}_0 \in U$ es un **extremo local o relativo**, si es mínimo local o máximo local. Un punto \vec{x}_0 en el interior del dominio de definición de la función es un **punto crítico** de f si $Df(\vec{x}_0) = \vec{0}$ o si el gradiente no existe. En el caso de gradiente nulo se dice también **punto estacionario**. Un punto estacionario que no es un extremo local se llama **punto silla**.

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

En el caso de una variable, hemos visto que si f tiene un extremo local en x_0 entonces o bien $f'(x_0) = 0$ o bien $f'(x_0)$ no existe. En el caso de varias variables, tenemos un resultado similar.

Teorema 11.2 Si $U \subset \mathbb{R}^n$ es abierto, $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable y $\vec{x}_0 \in U$ es un extremo local, entonces $Df(\vec{x}_0) = \vec{0}$; esto es, \vec{x}_0 es un punto crítico de f .

Demostración:

Consideremos sólo el caso en que f alcanza un máximo local, dado que el caso de mínimo local es completamente análogo. Entonces, para cualquier $\vec{h} \in \mathbb{R}^n$, la función $g(t) = f(\vec{x} + t\vec{h})$ tiene un máximo local en $t = 0$. Luego, por el cálculo de una variable $g'(0) = 0$, y usando la regla de la cadena se sigue que

$$g'(0) = [Df(\vec{x}_0)] \vec{h} = 0, \quad \forall \vec{h} \in \mathbb{R}^n \iff Df(\vec{x}_0) = \vec{0}.$$

c.q.d.

Si recordamos que $Df(\vec{x}_0) = \vec{0}$ significa que todas las componentes de $Df(\vec{x}_0)$ son "cero" podemos describir el resultado del teorema 11.2 diciendo que \vec{x}_0 es un punto crítico si cada derivada parcial es nula en \vec{x}_0 . En otras palabras si el gradiente en \vec{x}_0 es nulo:

$$\vec{\nabla} f(\vec{x}_0) = \vec{0}.$$

Estas ideas son muy generales y su aplicación a funciones de más de dos variables suele ser laboriosa. La mayoría de las veces nos limitaremos al caso de dos variables, pues además de ser los cálculos menos complicados, podemos hacer uso de nuestra intuición geométrica. Supongamos que $f = f(x, y)$ está definida en un conjunto abierto y es continuamente diferenciable. La gráfica de f es una superficie $z = f(x, y)$. Donde f presente un máximo local la superficie tendrá un punto localmente de máxima altura. Donde f presente un mínimo local la superficie tendrá un punto localmente de mínima altura. Donde f tiene o bien un máximo local o un mínimo local, su gradiente es $\vec{0}$, luego el plano tangente es horizontal. Un gradiente nulo señala la posible existencia de extremos locales, pero no la garantiza. Por ejemplo, en el caso de superficies en forma de silla (como el paraboloides hiperbólico que hemos encontrado en la sección 10.1), hay un plano tangente en el origen, luego el gradiente es cero en este punto, aunque en el origen no haya ni máximo ni mínimo local.

Ejemplo 1. Halla los máximos y mínimos de la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \rightsquigarrow x^2 + y^2$.

Solución:

Debemos identificar los puntos críticos de f resolviendo las ecuaciones

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

En este caso

$$\frac{\partial f}{\partial x}(x, y) = 2x, \quad \frac{\partial f}{\partial y}(x, y) = 2y.$$

CAPÍTULO 11. DERIVADAS DE ORDEN SUPERIOR: MÁXIMOS Y MÍNIMOS.

de manera que el único punto crítico es el origen $(0, 0)$, donde el valor de la función es nulo. Como $f(x, y) = x^2 + y^2 \geq 0$, este punto es un mínimo relativo. De hecho es un mínimo absoluto (o global) de la función f .

Ejemplo 2. Halla los máximos y mínimos de la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $(x, y) \rightsquigarrow x^2 - y^2$.

Solución:

Como en el ejemplo anterior, el único punto crítico es el origen $(0, 0)$, donde el valor de la función es nulo. Sin embargo cerca del origen obtenemos $f(x, 0) = x^2 \geq f(0, 0) = 0$ y $f(0, y) = -y^2 \leq f(0, 0) = 0$. Como se pueden tomar valores x o y arbitrariamente pequeños, el origen no puede ser ni mínimo ni máximo relativo, será un punto silla. Por lo tanto esta función no tiene extremos relativos.

Ejemplo 3. Halla los máximos y mínimos de la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por $f(x, y) = 2x^2 + y^2 - xy - 7y$.

Solución:

Debemos identificar los puntos críticos de f resolviendo las ecuaciones

$$\frac{\partial f}{\partial x}(x, y) = 4x - y = 0, \quad \frac{\partial f}{\partial y}(x, y) = 2y - x - 7 = 0.$$

La única solución del sistema es $x = 1$ e $y = 4$; por consiguiente el único punto crítico es $(1, 4)$. Comparemos ahora el valor de $f(1, 4)$ con los valores de f en los puntos próximos $(1 + h, 4 + k)$, hallando la diferencia $f(1 + h, 4 + k) - f(1, 4)$

$$\begin{aligned} f(1, 4) &= 2 + 16 - 4 - 28 = -14, \\ f(1 + h, 4 + k) &= 2(1 + h)^2 + (4 + k)^2 - (1 + h)(4 + k) - 7(4 + k) = \\ &= 2h^2 + k^2 - hk - 14, \quad \Rightarrow \\ f(1 + h, 4 + k) - f(1, 4) &= 2h^2 + k^2 - hk = \frac{1}{4}h^2 - hk + k^2 + \frac{7}{4}h^2 \\ &= \left(\frac{1}{2}h - k\right)^2 + \frac{7}{4}h^2 \end{aligned}$$

que es una cantidad no negativa para todo h y k reales. Se deduce entonces que f tiene un mínimo local en $(1, 4)$.

Ejemplo 4. Halla los máximos y mínimos de la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, definida por $f(x, y) = y^2 - xy + 2x + y + 1$.

Solución:

Debemos identificar los puntos críticos de f resolviendo las ecuaciones

$$\frac{\partial f}{\partial x}(x, y) = -y + 2 = 0, \quad \frac{\partial f}{\partial y}(x, y) = 2y - x + 1 = 0.$$

La única solución del sistema es $x = 5$ e $y = 2$; por consiguiente el único punto crítico es $(5, 2)$. Comparemos ahora el valor de $f(5, 2)$ con los valores de f en los puntos próximos

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

$(5 + h, 2 + k)$, hallando la diferencia $f(5 + h, 2 + k) - f(5, 2)$

$$\begin{aligned} f(5, 2) &= 4 - 10 + 10 + 2 + 1 = 7, \\ f(5 + h, 2 + k) &= (2 + k)^2 - (5 + h)(2 + k) + 2(5 + h) + (2 + k) + 1 = \\ &= k^2 - hk + 7, \quad \Rightarrow \\ f(5 + h, 2 + k) - f(5, 2) &= k^2 - hk = k(k - h) \end{aligned}$$

que es una cantidad que no tiene signo constante para h y k pequeños. Se deduce entonces que f tiene un punto silla en $(5, 2)$.

En todos estos ejemplos, las funciones consideradas eran diferenciables en todo su dominio de definición. Los únicos puntos críticos eran puntos estacionarios. Consideremos ahora una función definida en todo punto pero que no es diferenciable en el origen. Resulta que el origen es un punto crítico aunque no un punto estacionario. Un ejemplo viene dado por la función

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

que está definida en todo \mathbb{R}^2 . Su gráfica es la hoja superior de un cono, y evidentemente el valor $f(0, 0) = 1$ es un mínimo local. Sin embargo, las derivadas parciales

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

no están definidas en $(0, 0)$, luego el gradiente no está definido en este punto y por lo tanto el punto $(0, 0)$ es un punto crítico aunque no un punto estacionario. El punto $(0, 0, 1)$ es un pico de la superficie y no hay plano tangente.

Pasamos ahora a deducir un criterio que dependa de la segunda derivada, para que un punto crítico sea extremo relativo. Este criterio se reducirá a la condición conocida de que $f''(x) > 0$ para un mínimo y $f''(x) < 0$ para un máximo.

Demos previamente unas definiciones.

Definición. Las **funciones cuadráticas** son funciones $g: \mathbb{R}^n \rightarrow \mathbb{R}$ que tienen la forma:

$$g(h_1, \dots, h_n) = \sum_{i,j=1}^n a_{ij} h_i h_j$$

para una matriz $[a_{ij}]$ (que podemos suponer simétrica), que en términos de multiplicación de matrices se puede escribir como

$$g(h_1, \dots, h_n) = (h_1, \dots, h_n) \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

La función cuadrática g será **definida positiva** si $g(\vec{h}) \geq 0$, para todo $\vec{h} \in \mathbb{R}^n$ y $g(\vec{h}) = 0$ sólo para $\vec{h} = \vec{0}$; de manera análoga, g será **definida negativa** si $g(\vec{h}) \leq 0$, para todo $\vec{h} \in \mathbb{R}^n$ y $g(\vec{h}) = 0$ sólo para $\vec{h} = \vec{0}$.

Definición. Sea $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ con derivadas parciales de segundo orden $\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}_0)$ para $i, j = 1, \dots, n$ en un punto $\bar{x}_0 \in U$. El **hessiano de f en \bar{x}_0** es la función cuadrática $Hf(\bar{x}_0): \mathbb{R}^n \rightarrow \mathbb{R}$, definida por

$$Hf(\bar{x}_0)(\vec{h}) = \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}_0) h_i h_j.$$

Nótese que si $n = 1$, el hessiano se reduce a $Hf(x_0)(h) = \frac{1}{2} f''(x_0) h^2$, que es definido positivo si $f''(x_0) > 0$. Podemos entonces enunciar y demostrar el criterio para extremos relativos.

Teorema 11.3 Si $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase C^3 , $\bar{x}_0 \in U$ es un punto crítico de f y el hessiano $Hf(\bar{x}_0)$ es definido positivo, entonces \bar{x}_0 es un mínimo relativo de f . De manera análoga, si el hessiano $Hf(\bar{x}_0)$ es definido negativo, entonces \bar{x}_0 es un máximo relativo de f .

Demostración:

La demostración requiere el teorema de Taylor (teorema 11.1) y un resultado de álgebra lineal (teorema 11.4) que se da a continuación.

Si $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ es de clase C^3 y $\bar{x}_0 \in U$ es un punto crítico de f , el teorema de Taylor se puede expresar en la forma

$$f(\bar{x}_0 + \vec{h}) - f(\bar{x}_0) = Hf(\bar{x}_0)(\vec{h}) + R_2(\vec{h}, \bar{x}_0)$$

donde $\frac{R_2(\vec{h}, \bar{x}_0)}{\|\vec{h}\|^2} \rightarrow 0$ cuando $\vec{h} \rightarrow \vec{0}$. Como $Hf(\bar{x}_0)$ es definido positivo, el teorema 11.4 asegura la existencia de una constante $M > 0$ tal que para todo $\vec{h} \in \mathbb{R}^n$

$$Hf(\bar{x}_0)(\vec{h}) \geq M \|\vec{h}\|^2.$$

Como $\frac{R_2(\vec{h}, \bar{x}_0)}{\|\vec{h}\|^2} \rightarrow 0$ cuando $\vec{h} \rightarrow \vec{0}$, existen $\hat{M} > 0$ y $\delta > 0$ tal que para $0 < \|\vec{h}\| < \delta$

$$\left| R_2(\vec{h}, \bar{x}_0) \right| > \hat{M} \|\vec{h}\|^2.$$

Luego, para $0 < \|\vec{h}\| < \delta$

$$0 < Hf(\bar{x}_0)(\vec{h}) + R_2(\vec{h}, \bar{x}_0) = f(\bar{x}_0 + \vec{h}) - f(\bar{x}_0)$$

es decir, \bar{x}_0 es un mínimo relativo estricto.

El caso definido negativo es análogo.

c.q.d.

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

Teorema 11.4 Si $B = [b_{ij}]$ es una matriz real $n \times n$ y si la función cuadrática asociada

$$H: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (h_1, \dots, h_n) \rightsquigarrow \frac{1}{2} \sum_{i,j=1}^n b_{ij} h_i h_j$$

es definida positiva, entonces existe una constante $M > 0$ tal que para todo $\vec{h} \in \mathbb{R}^n$

$$H(\vec{h}) \geq M \|\vec{h}\|^2.$$

Demostración:

Para $\|\vec{h}\| = 1$ definimos la función $g(\vec{h}) = H(\vec{h})$, que es una función continua de \vec{h} , y por lo tanto alcanza un valor mínimo, digamos M . Como H es cuadrática, tenemos

$$H(\vec{h}) = H\left(\frac{\vec{h}}{\|\vec{h}\|} \|\vec{h}\|\right) = H\left(\frac{\vec{h}}{\|\vec{h}\|}\right) \|\vec{h}\|^2 = g\left(\frac{\vec{h}}{\|\vec{h}\|}\right) \|\vec{h}\|^2 \geq M \|\vec{h}\|^2,$$

para cualquier $\vec{h} \neq \vec{0}$.

c.q.d.

Antes de pasar a ver algún ejemplo, presentamos un criterio útil para saber cuándo una función cuadrática definida por una matriz 2×2 es definida positiva.

Criterio.

Sea $B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ una matriz 2×2 y $H(\vec{h}) = \frac{1}{2} (h_1, h_2) B \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$, la función cuadrática asociada a la matriz B . Entonces, $H(\vec{h})$ es definida positiva si y sólo si

$$a > 0 \quad \text{y} \quad \det B = ac - b^2 > 0.$$

De manera análoga, $H(\vec{h})$ es definida negativa si y sólo si

$$a < 0 \quad \text{y} \quad \det B = ac - b^2 > 0.$$

Hay criterios similares para una matriz simétrica $n \times n$: considerando las submatrices diagonales, si todos los determinantes de estas submatrices son positivos, tendremos que la función cuadrática asociada con la matriz original es definida positiva. Si los signos se alternan en el orden negativo-positivo, será definida negativa.

Este criterio y el teorema 11.3 dan lugar al resultado siguiente.

Teorema 11.5 Sea $f(x, y)$ de clase C^3 en un conjunto abierto $U \subset \mathbb{R}^2$. Un punto (x_0, y_0) es un mínimo local (estricto) de f si se cumplen las tres condiciones siguientes:

$$(i) \quad \frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0;$$

$$(ii) \quad \frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0;$$

$$(iii) \quad D = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial y^2} \end{pmatrix} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 > 0 \text{ en } (x_0, y_0).$$

Si en (ii) tenemos < 0 en lugar de > 0 sin cambiar la condición (iii) del discriminante D , entonces tenemos un máximo local (estricto).

Si se tuviera $D < 0$, entonces tendríamos un punto silla.

Ejemplo 1. Clasifica los puntos críticos de la función $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, dada por $(x, y) \rightsquigarrow x^2 - 2xy + 2y^2$.

Solución:

Como en el caso de los ejemplos anteriores, debemos identificar los puntos críticos de f resolviendo las ecuaciones

$$\frac{\partial f}{\partial x}(x, y) = 2x - 2y = 0, \quad \frac{\partial f}{\partial y}(x, y) = -2x + 4y = 0.$$

La única solución del sistema es $x = 0$ e $y = 0$. El hessiano resulta ser

$$Hf(\vec{0}) \left(\vec{h} \right) = h_1^2 - 2h_1 h_2 + 2h_2^2 = (h_1 - h_2)^2 + h_2^2$$

es cual es claramente definido positivo. Así, f tiene un mínimo relativo en $(0, 0)$. De manera alternativa, podemos aplicar el teorema 11.5: en el punto $(0, 0)$ tenemos

$$\frac{\partial^2 f}{\partial x^2}(0, 0) = 2, \quad \frac{\partial^2 f}{\partial y^2}(0, 0) = 4 \quad \text{y} \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = -2$$

es decir se cumplen las tres condiciones

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

$$(ii) \quad \frac{\partial^2 f}{\partial x^2}(0, 0) = 2 > 0,$$

$$(iii) \quad D = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} \\ \frac{\partial^2 f}{\partial y^2} \end{pmatrix} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 = 2 \cdot 4 - (-2)^2 = 4 > 0 \text{ en } (0, 0).$$

Ejemplo 2. La gráfica de la función $g(x, y) = \frac{1}{xy}$ es una superficie S en \mathbb{R}^3 . Halla los puntos en S más cercanos al origen $(0, 0, 0)$.

Solución:

La distancia de (x, y, z) a $(0, 0, 0)$ está dada por la fórmula $d(x, y, z) = \sqrt{x^2 + y^2 + z^2}$.

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

Si el punto (x, y, z) está en la superficie S , entonces se puede expresar la distancia como una función de dos variables

$$d^*(x, y) = d\left(x, y, \frac{1}{xy}\right) = \sqrt{x^2 + y^2 + \frac{1}{x^2 y^2}}.$$

Como $d^* > 0$, esta distancia se minimizará cuando la nueva función $f(x, y) = [d^*(x, y)]^2$ sea mínima. Calculemos entonces el gradiente de la función

$$f(x, y) = [d^*(x, y)]^2 = x^2 + y^2 + \frac{1}{x^2 y^2};$$

será

$$\vec{\nabla} f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \left(2x - \frac{2}{x^3 y^2}, 2y - \frac{2}{x^2 y^3}\right),$$

que es nulo si y sólo si

$$2x - \frac{2}{x^3 y^2} = 0 = 2y - \frac{2}{x^2 y^3} \Leftrightarrow \begin{cases} x^4 y^2 - 1 = 0 \\ x^2 y^4 - 1 = 0 \end{cases} \Leftrightarrow \begin{cases} x = \pm 1 \\ y = \pm 1 \end{cases}$$

Se sigue entonces que f tiene cuatro puntos críticos: $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$.

Para determinar en qué casos se trata de mínimos locales, máximos locales o puntos silla, aplicamos el teorema 11.5:

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 2 + \frac{6}{x^4 y^2}, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 2 + \frac{6}{x^2 y^4} \quad \text{y} \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{4}{x^3 y^3}.$$

En los cuatro puntos críticos resulta que

$$\frac{\partial^2 f}{\partial x^2}(1, 1) = \frac{\partial^2 f}{\partial y^2}(1, 1) = 8, \quad \frac{\partial^2 f}{\partial x \partial y}(1, 1) = 4 \Rightarrow D = 64 - 16 = 48 > 0,$$

$$\frac{\partial^2 f}{\partial x^2}(1, -1) = \frac{\partial^2 f}{\partial y^2}(1, -1) = 8, \quad \frac{\partial^2 f}{\partial x \partial y}(1, -1) = -4 \Rightarrow D = 48 > 0,$$

$$\frac{\partial^2 f}{\partial x^2}(-1, 1) = \frac{\partial^2 f}{\partial y^2}(-1, 1) = 8, \quad \frac{\partial^2 f}{\partial x \partial y}(-1, 1) = -4 \Rightarrow D = 48 > 0,$$

$$\frac{\partial^2 f}{\partial x^2}(-1, -1) = \frac{\partial^2 f}{\partial y^2}(-1, -1) = 8, \quad \frac{\partial^2 f}{\partial x \partial y}(-1, -1) = 4 \Rightarrow D = 48 > 0,$$

de modo que cada uno de los puntos críticos es mínimo local de f y estos son todos los mínimos locales de la función f . En estos puntos la función $f(x, y)$ asume valor igual a 3. de modo que los puntos sobre la superficie más cercanos al origen $(0, 0, 0)$ son $(1, 1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$, $(-1, -1, 1)$, con distancia $d^* = \sqrt{3}$.

Ejemplo 3. Analiza el comportamiento de $z = x^5 y + x y^5 + x y$ en sus puntos críticos.

Solución:

Las primeras derivadas parciales son

$$\frac{\partial z}{\partial x} = 5x^4 y + y^5 + y = y(5x^4 + y^4 + 1), \quad \frac{\partial z}{\partial y} = x^5 + 5xy^4 + x = x(x^4 + 5y^4 + 1).$$

dado que los términos entre paréntesis son siempre mayores o iguales que 1, el único punto crítico está en $(0, 0)$.

Las segundas derivadas parciales son

$$\frac{\partial^2 z}{\partial x^2}(x, y) = 20x^3 y, \quad \frac{\partial^2 z}{\partial y^2}(x, y) = 20xy^3 \quad \text{y} \quad \frac{\partial^2 z}{\partial x \partial y}(x, y) = 5x^4 + 5y^4 + 1$$

que evaluadas en $(0, 0)$ nos dan

$$\frac{\partial^2 z}{\partial x^2}(0, 0) = 0, \quad \frac{\partial^2 z}{\partial y^2}(0, 0) = 0, \quad \frac{\partial^2 z}{\partial x \partial y}(0, 0) = 1 \quad \text{y} \quad D = -1 < 0.$$

Es decir el punto crítico $(0, 0)$ es un punto silla.

Así como en el caso de funciones de una variable, el criterio de la derivada segunda se aplica a puntos x_0 en los cuales $g'(x_0) = 0$ pero $g''(x_0) \neq 0$. Si $g''(x_0) = 0$, el criterio de la derivada segunda no proporciona ninguna información. El criterio de la derivada parcial segunda del teorema 11.5 tiene la misma limitación. Vale para los puntos (x_0, y_0) tales que $\vec{\nabla} f(x_0, y_0) = \vec{0}$, pero $D \neq 0$. Si $D = 0$, el criterio de la derivada segunda no proporciona información.

Consideremos, por ejemplo, las funciones

$$F(x, y) = x^4 + y^4, \quad G(x, y) = -x^4 - y^4, \quad H(x, y) = x^4 - y^4.$$

Cada una de esas funciones tiene un gradiente nulo en el origen y en cada caso es $D = 0$. No obstante, claramente para la función $F(x, y)$ el origen $(0, 0)$ es un punto de mínimo local y para la función $G(x, y)$ el origen $(0, 0)$ es un punto de máximo local. Por otro lado para la función $H(x, y)$ el origen $(0, 0)$ es un punto de silla dado que $H(0, 0) = 0$, pero en cualquier entorno de $(0, 0)$, la función H toma tanto valores positivos como valores negativos:

$$H(x, 0) > 0 \quad \text{para} \quad x \neq 0 \quad \text{mientras que} \quad H(0, y) < 0 \quad \text{para} \quad y \neq 0.$$

Definición. Sea $f: A \rightarrow \mathbb{R}$ una función definida en un subconjunto A de \mathbb{R}^2 ó \mathbb{R}^3 . Se dice que un punto $\vec{x}_0 \in A$ es un punto de **máximo absoluto** de f si $f(\vec{x}) \leq f(\vec{x}_0)$ para todo $\vec{x} \in A$. De manera similar se dice que un punto $\vec{x}_0 \in A$ es un punto de **mínimo absoluto** de f si $f(\vec{x}) \geq f(\vec{x}_0)$ para todo $\vec{x} \in A$.

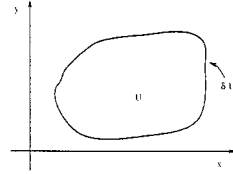
En el caso de una función f de una variable real continua en un intervalo cerrado (y entonces acotado), sabemos que (ver teorema 0.11) $f(x)$ alcanza un máximo absoluto y

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

un mínimo absoluto en dicho intervalo. También se cumple una generalización a \mathbb{R}^n de este hecho.

Recordemos antes unas definiciones ya dadas en la sección de límites y continuidad, 10.2.

Definición. Se dice que un conjunto $D \subset \mathbb{R}^n$ es **acotado** si existe un número $M > 0$ tal que $\|\vec{x}\| < M$ para todo $\vec{x} \in D$. Un conjunto es **cerrado** si contiene a todos sus puntos de frontera, que denotamos con ∂D . Así, un conjunto es acotado si puede estar estrictamente contenido en una bola. En la figura se muestra el conjunto $D = U \cup \partial U$. Claramente $\partial D = \partial U$.



Teorema 11.6 Teorema del máximo y del mínimo

Sea D cerrado y acotado en \mathbb{R}^n y sea $f: D \rightarrow \mathbb{R}$ continua. Entonces, f alcanza sus valores máximo y mínimo en algunos puntos \vec{x}_0 y \vec{x}_1 de D .

Lo que este teorema nos dice es que \vec{x}_0 y \vec{x}_1 son puntos en D donde f toma el mayor o el menor valor posible. Como en el caso de una variable, estos puntos no necesariamente están determinados de manera única. Una consecuencia del teorema 11.2 es la siguiente

Teorema 11.7 Sea D cerrado y acotado en \mathbb{R}^n , es decir $D = U \cup \partial U$, con U abierto y ∂U su frontera. Sea $f: D \rightarrow \mathbb{R}$ una función continua, y además de clase C^1 en U . Si f alcanza su valor máximo (o mínimo) en un punto \vec{x}_0 de U , entonces \vec{x}_0 es un punto crítico de f .

Para hallar el máximo y mínimo absolutos de una función de clase C^1 , $f: D \rightarrow \mathbb{R}$, empleamos un procedimiento similar al del cálculo de una variable:

- (i) localizamos todos los puntos críticos de f en U ;
- (ii) hallamos los puntos críticos de f considerada como función definida sólo en ∂U ;
- (iii) calculamos el valor de f en todos estos puntos críticos;
- (iv) comparamos estos valores y seleccionamos el mayor y el menor.

Para acabar esta sección presentamos unos ejemplos sencillos de cómo manejar el paso (ii); en una sección siguiente (sección 11.3) se presentará otro método, el método de los multiplicadores de Lagrange.

Ejemplo 1. Encuentra los máximos y mínimos absolutos en el dominio triangular dado por $D = \{(x, y) / 0 \leq x \leq 4, 0 \leq y \leq 2x\}$ para la función $f(x, y) = xy - 2x - 3y$.

Solución:

Por el teorema 11.6 sabemos que f alcanza en D un máximo absoluto y un mínimo absoluto. Para ver si algunos de esos valores se alcanza en el interior, consideremos el gradiente

$$\vec{\nabla} f(x, y) = (y - 2, x - 3) \quad \Rightarrow \quad \vec{\nabla} f(x, y) = \vec{0} \quad \Leftrightarrow \quad \begin{cases} x = 3 \\ y = 2 \end{cases}$$

Las segundas derivadas parciales son

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

Visto que $D = -1 < 0$, resulta que el punto crítico $(3, 2)$ es un punto silla.

Vamos a buscar ahora los valores extremos en la frontera, escribiendo cada lado del triángulo en la forma $\vec{r} = \vec{r}(t)$ y analizando $f(\vec{r}(t))$. Con los segmentos C_1, C_2, C_3 dados en la figura tenemos

$$C_1: y = 2x \quad \text{con} \quad 0 \leq x \leq 4$$

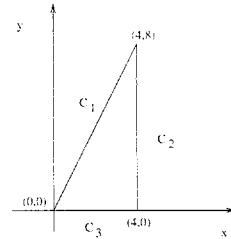
$$\vec{r}_1(t) = t\vec{i} + 2t\vec{j}, \quad t \in [0, 4]$$

$$C_2: x = 4 \quad \text{con} \quad 0 \leq y \leq 8$$

$$\vec{r}_2(t) = 4\vec{i} + t\vec{j}, \quad t \in [0, 8]$$

$$C_3: y = 0 \quad \text{con} \quad 0 \leq x \leq 4$$

$$\vec{r}_3(t) = t\vec{i}, \quad t \in [0, 4]$$



Los valores de f en estos segmentos de recta vienen dados por las funciones

$$f_1(t) = 2t^2 - 8t, \quad t \in [0, 4]; \quad f_2(t) = t - 8, \quad t \in [0, 8]; \quad f_3(t) = -2t, \quad t \in [0, 4];$$

cuyas derivadas son

$$f'_1(t) = 4t - 8, \quad t \in [0, 4]; \quad f'_2(t) = 1, \quad t \in [0, 8]; \quad f'_3(t) = -2, \quad t \in [0, 4].$$

La primera derivada $f'_1(t)$ se anula para $t = 2$, y dado que $f''_1(t) = 4 > 0$ éste será un punto de mínimo local para f en el segmento de recta. Su valor es $f_1(2) = -8$.

La derivada de la función $f_2(t)$ resulta ser siempre positiva, lo que quiere decir que la $f_2(t)$ es una función estrictamente creciente, es decir su mínimo absoluto se alcanza en el extremo izquierdo del intervalo $[0, 8]$, $f_2(0) = -8$, mientras que el máximo absoluto se alcanza en el extremo derecho del intervalo $[0, 8]$, $f_2(8) = 0$.

La derivada de la función $f_3(t)$ resulta ser siempre negativa, lo que quiere decir que la $f_3(t)$ es una función estrictamente decreciente, es decir su máximo absoluto se alcanza en el extremo izquierdo del intervalo $[0, 4]$, $f_3(0) = 0$, mientras que el mínimo absoluto se alcanza en el extremo derecho del intervalo $[0, 4]$, $f_3(4) = -8$.

De ello se deduce que el valor máximo de f es 0, mientras que el valor mínimo es -8 . Además, $f(0, 0) = f(4, 8) = 0$, $f(4, 0) = f(2, 4) = -8$.

11.2. EXTREMOS DE FUNCIONES CON VALORES REALES.

Ejemplo 2. El rectángulo $R_1 = \{(x, y) / 0 \leq x \leq a, -b \leq y \leq b\}$ es un subconjunto cerrado acotado del plano. La función

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

al ser continua en todo punto, es continua en el rectángulo. Luego podemos estar seguros de que f alcanza un máximo absoluto y un mínimo absoluto en ese conjunto. El máximo absoluto se alcanza en los puntos $(a, -b)$ y (a, b) (¿por qué?), y el valor de la función en estos puntos es $1 + \sqrt{a^2 + b^2}$. El mínimo absoluto se alcanza en el origen (¿por qué?); aquí el valor es 1.

Si ahora consideramos el rectángulo $R_2 = \{(x, y) / 0 < x \leq a, -b \leq y \leq b\}$, éste está acotado pero no es cerrado. En este conjunto la misma función f alcanza el mismo valor máximo y en los mismos puntos de antes, pero no alcanza un valor mínimo, dado que el origen no está en el conjunto.

Finalmente, si consideramos la misma función en todo el plano, que es un conjunto cerrado pero no acotado, f alcanza un mínimo absoluto en el origen, pero no un máximo absoluto.

Ejercicios.

1) Halla los puntos críticos de las funciones siguientes y determina cuáles son máximos locales, mínimos locales o puntos silla.

a) $f(x, y) = x^2 - y^2 + xy;$

b) $f(x, y) = x^2 + y^2 - xy;$

c) $f(x, y) = x^2 + y^2 + 2xy;$

d) $f(x, y) = x^2 + y^2 + 3xy;$

e) $f(x, y) = x^2 - 3xy + 5x - 2y + 6y^2 + 8;$

f) $f(x, y) = e^{1+x^2-y^2};$

g) $f(x, y) = y + x \operatorname{sen} y;$

h) $f(x, y) = e^x \cos y;$

e) $f(x, y) = (x+y)(xy+1);$

d) $f(x, y) = (x-y)(xy-1).$

2) Sea $f(x, y) = x^2 - 2xy + y^2$. Aquí $D = 0$. ¿Qué puntos críticos son mínimos locales, máximos locales o puntos silla?

3) Dada la función $f(x, y, z) = x^2 + y^2 + z^2 + xy$, determina la naturaleza de los puntos críticos.

4) Halla los puntos críticos de las funciones siguientes y determina los valores extremos locales.

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

b) $f(x, y) = x^2 + 2y^2 - 4y;$

c) $f(x, y) = x^3 - 3x + y;$

d) $f(x, y) = 4x^3y - 4xy^3.$

5) Halla los valores extremos absolutos tomados por la función en el conjunto D .

a) $f(x, y) = \frac{1}{\sqrt{x^2 + y^2}}, \quad D = \{(x, y) / 1 \leq x \leq 3, 1 \leq y \leq 4\};$

b) $f(x, y) = 4x^2 - 9y^2, \quad D = \{(x, y) / -1 \leq x \leq 1\};$

- c) $f(x, y) = (x - y)^2$, $D = \{(x, y) / 0 \leq x \leq 6, 0 \leq y \leq 12 - 2x\}$;
 d) $f(x, y) = (x - 3)^2 + y^2$, $D = \{(x, y) / 0 \leq x \leq 4, x^2 \leq y \leq 4x\}$;
 d) $f(x, y) = xy(1 - x^2 - y^2)$, $D = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq 1\}$.

11.3 Extremos con restricciones y multiplicadores de Lagrange.

Cuando preguntamos por la distancia que hay desde el punto $P(x_0, y_0)$ a la recta l de ecuación $Ax + By + C = 0$, estamos preguntando por el valor mínimo de la función

$$f(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2},$$

donde (x, y) ha de satisfacer la condición suplementaria, llamada ligadura o restricción. $Ax + By + C = 0$. De la misma manera, cuando preguntamos por la distancia que hay desde el punto $P(x_0, y_0, z_0)$ al plano de ecuación $Ax + By + Cz + D = 0$, estamos preguntando por el valor mínimo de la función

$$f(x, y, z) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2},$$

donde (x, y, z) ha de satisfacer la restricción $Ax + By + Cz + D = 0$.

Ya hemos tratado estos problemas particulares mediante alguna técnica especial. Nuestro interés aquí es presentar técnicas que permitan tratar problemas más generales. En el caso de dos variables, el problema consistirá en maximizar o minimizar una expresión $f(x, y)$ sujeta a una restricción $g(x, y) = 0$. En el caso de tres variables, el problema consistirá en maximizar o minimizar una expresión $f(x, y, z)$ sujeta a una restricción $g(x, y, z) = 0$.

Empecemos con un **ejemplo**. Queremos hallar el volumen máximo de un paralelepípedo rectangular sabiendo que la suma de las longitudes de las aristas es $12a$.

Sean x, y, z las longitudes de las aristas del paralelepípedo. El volumen viene dado por

$$V = xyz$$

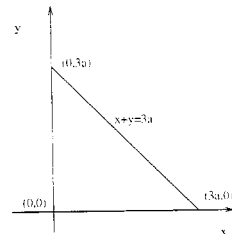
y la restricción sobre las aristas es

$$4(x + y + z) = 12a.$$

Despejando la z de la última ecuación hallamos $z = 3a - (x + y)$, y sustituyendo en la expresión del volumen obtenemos la función de dos variables

$$V = xy [3a - (x + y)].$$

Dado que $x, y, z = 3a - (x + y)$ han de ser positivos, nuestro problema consiste en hallar el valor máximo de V en el interior del triángulo mostrado en la figura.



11.3. EXTREMOS CON RESTRICCIONES Y MULTIPLICADORES DE LAGRANGE.

Las derivadas parciales de primer orden son

$$\frac{\partial V}{\partial x} = 3ay - 2xy - y^2 \quad \frac{\partial V}{\partial y} = 3ax - 2xy - x^2.$$

Igualando ambas derivadas a cero, obtenemos

$$y(3a - 2x - y) = 0, \quad x(3a - 2y - x) = 0.$$

Dado que x e y han de ser positivos, podemos dividir por x y por y obteniendo,

$$3a - 2x - y = 0, \quad 3a - 2y - x = 0.$$

Resolviendo el sistema formado por estas ecuaciones, hallamos que $x = y = a$. El punto (a, a) que está dentro del triángulo es el único punto estacionario. El valor del volumen V en dicho punto es a^3 . Las condiciones del problema hacen evidente que este valor es un máximo (confirmarlo usando el criterio de la derivada segunda).

Este ejemplo es fácil, en parte debido a que la restricción se podía resolver expresando una de las variables en función de las otras. En general, esto no es posible y por esto se precisa de un método más sofisticado.

Para que las cosas sean más sencillas consideremos sólo el caso de dos y tres variables, es decir a lo largo de la sección f será una función de dos o tres variables de valores reales continuamente diferenciable en algún conjunto abierto U . Sea C una curva enteramente contenida en U , $\vec{\sigma} = \vec{\sigma}(t)$, con $t \in I \subset \mathbb{R}$ y que tiene en cada punto un vector tangente no nulo $\vec{\sigma}'(t)$. Un resultado básico es el siguiente.

Teorema 11.8 Si \vec{x}_0 maximiza o minimiza la función $f(\vec{x})$ en C , $\vec{\nabla} f(\vec{x}_0)$ es perpendicular a C en el punto \vec{x}_0 .

Demostración:

Escojamos t_0 tal que $\vec{\sigma}(t_0) = \vec{x}_0$. La función compuesta $f(\vec{\sigma}(t))$ tiene un máximo o un mínimo en t_0 . Consecuentemente su derivada $\frac{d}{dt} [f(\vec{\sigma}(t))] = \vec{\nabla} f(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t)$ ha de ser nula en t_0 :

$$\frac{d}{dt} [f(\vec{\sigma}(t_0))] = \vec{\nabla} f(\vec{\sigma}(t_0)) \cdot \vec{\sigma}'(t_0) = \vec{\nabla} f(\vec{x}_0) \cdot \vec{\sigma}'(t_0) = 0.$$

Esto demuestra que $\vec{\nabla} f(\vec{x}_0)$ es ortogonal a $\vec{\sigma}'(t_0)$. Dado que $\vec{\sigma}'(t_0)$ es tangente a la curva C en el punto \vec{x}_0 , se sigue que $\vec{\nabla} f(\vec{x}_0)$ es perpendicular a C en \vec{x}_0 .

c.q.d

Estamos ahora en condiciones de considerar el problema con restricciones. Supondremos que como f , también g es una función de dos o tres variables de valores reales continuamente diferenciable definida en un subconjunto del dominio de f . Lagrange hizo la siguiente observación general.

Teorema 11.9 Teorema de los multiplicadores de Lagrange.

Sean $f: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $g: U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funciones suaves. Sean $\vec{x}_0 \in U$ y $g(\vec{x}_0) = c$, y sea $S = \{\vec{x} \in \mathbb{R}^n / g(\vec{x}) = c\}$ el conjunto de nivel para g con valor c .

Suponer que $\vec{\nabla} g(\vec{x}_0) \neq \vec{0}$.

Si $f|_S$, que denota a "f restringida a S", tiene un máximo o un mínimo en S en el punto \vec{x}_0 , entonces existe un número λ , que se llama multiplicador de Lagrange, tal que

$$\vec{\nabla} f(\vec{x}_0) = \lambda \vec{\nabla} g(\vec{x}_0). \quad (11.5)$$

Demostración:

No somos capaces de realizar una demostración completa del caso general; sin embargo lo podemos hacer en los casos $n = 2, 3$.

Supongamos que \vec{x}_0 maximiza (o minimiza) la función $f(\vec{x})$ sujeta a la restricción $g(\vec{x}) = 0$. Si $\vec{\nabla} f(\vec{x}_0) = \vec{0}$ el resultado es trivialmente cierto, cualquier vector es paralelo al vector nulo. Supongamos entonces que $\vec{\nabla} f(\vec{x}_0) \neq \vec{0}$.

En el caso de dos variables tenemos

$$\vec{x}_0 = (x_0, y_0) \quad \text{y la restricción} \quad g(x, y) = 0.$$

La restricción define una curva $S = C$ que tiene un vector tangente no nulo en (x_0, y_0) (ver la definición dada en la sección 10.5)

$$\vec{t}(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) \vec{i} - \frac{\partial f}{\partial x}(x_0, y_0) \vec{j} \neq \vec{0}.$$

Dado que (x_0, y_0) maximiza (o minimiza) $f(x, y)$ en C , sabemos por el teorema 11.8 que $\vec{\nabla} f(x_0, y_0)$ es perpendicular a C en el punto (x_0, y_0) . Luego los dos gradientes son paralelos.

En el caso de tres variables tenemos

$$\vec{x}_0 = (x_0, y_0, z_0) \quad \text{y la restricción} \quad g(x, y, z) = 0.$$

La restricción define una superficie $S = \Gamma$ que está en el dominio de f . Sea ahora C una curva que está en Γ y pasa por el punto (x_0, y_0, z_0) , siendo su vector tangente no nulo en (x_0, y_0, z_0) . Sabemos que (x_0, y_0, z_0) maximiza (o minimiza) $f(x, y, z)$ en C y en consecuencia (por el teorema 11.8) que $\vec{\nabla} f(x_0, y_0, z_0)$ es perpendicular a C en el punto (x_0, y_0, z_0) . Dado que esto es cierto para cualquier curva C , el gradiente $\vec{\nabla} f(x_0, y_0, z_0)$ ha de ser perpendicular a la superficie Γ . Pero como vimos en la sección 10.5, también $\vec{\nabla} g(x_0, y_0, z_0)$ es perpendicular a Γ en (x_0, y_0, z_0) . De ello se deduce que los dos gradientes $\vec{\nabla} f(x_0, y_0, z_0)$ y $\vec{\nabla} g(x_0, y_0, z_0)$ son paralelos, es decir

$$\vec{\nabla} f(\vec{x}_0) = \lambda \vec{\nabla} g(\vec{x}_0).$$

c.q.d

11.3. EXTREMOS CON RESTRICCIONES Y MULTIPLICADORES DE LAGRANGE.

Ejemplo 1. Sea $S \subset \mathbb{R}^2$ la recta que pasa por $(-1, 0)$ inclinada a 45° , y sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $(x, y) \rightsquigarrow x^2 + y^2$. Halla los extremos de $f|_S$.

Solución:

Aquí, $S = \{(x, y) / y - x - 1 = 0\}$, y por lo tanto la restricción será $g(x, y) = y - x - 1 = 0$ con $c = 0$. Tenemos que $\vec{\nabla} g(x, y) = -\vec{i} + \vec{j}$. Según el teorema 11.8 los extremos relativos de $f|_S$ deben hallarse entre los puntos en que $\vec{\nabla} f$ es ortogonal a S , esto es, inclinado a -45° . Pero $\vec{\nabla} f(x, y) = (2x, 2y)$, que tiene la pendiente deseada sólo cuando $x = -y$, o cuando (x, y) está sobre la recta L que pasa por el origen inclinada a -45° . Esto puede suceder en el conjunto S sólo para el único punto en el que se intersecan L y S , es decir $\left(-\frac{1}{2}, \frac{1}{2}\right)$, que resulta ser un punto de mínimo.

Ejemplo 2. Maximiza y minimiza en la circunferencia de radio unidad la función

$$f(x, y) = xy.$$

Solución:

Dado que f es continua y que la circunferencia $x^2 + y^2 = 1$ es cerrada, por el teorema 11.6 se alcanza el máximo y el mínimo.

Para aplicar el teorema de Lagrange, teorema 11.9, hacemos

$$g(x, y) = x^2 + y^2 - 1, \quad \text{con } c = 0.$$

Entonces queremos maximizar y minimizar

$$f(x, y) = xy \quad \text{con la restricción } g(x, y) = x^2 + y^2 - 1 = 0.$$

Los gradientes son

$$\vec{\nabla} f(x, y) = y\vec{i} + x\vec{j}, \quad \vec{\nabla} g(x, y) = 2x\vec{i} + 2y\vec{j}.$$

La condición (11.5), $\vec{\nabla} f(x_0) = \lambda \vec{\nabla} g(x_0)$ nos da dos ecuaciones, lo que junto a la ecuación que da la restricción, nos permite escribir un sistema (en general no lineal) de tres ecuaciones en las tres incógnitas x , y , λ :

$$\begin{cases} y = 2\lambda x \\ x = 2\lambda y \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} y^2 = 2\lambda xy \\ x^2 = 2\lambda xy \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} y^2 = x^2 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow 2x^2 = 1 \Rightarrow$$

$$x = \pm \frac{1}{\sqrt{2}} \quad \text{es decir} \quad \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right),$$

son los únicos puntos que pueden dar lugar a un valor extremo. En el primero y en el cuarto punto la función f toma el valor $\frac{1}{2}$; mientras que en los otros dos toma valor $-\frac{1}{2}$.

Entonces claramente, $\frac{1}{2}$ es el valor máximo y $-\frac{1}{2}$ es el valor mínimo.

Observación. Algunas recomendaciones generales son útiles para problemas de este tipo. En primer lugar si la curva C o la superficie S está acotada, entonces por el teorema 11.6 f debe tener un máximo y un mínimo. En particular si f sólo tiene dos puntos que satisfagan las condiciones del teorema de los multiplicadores de Lagrange o del teorema 11.8, entonces uno debe ser un máximo y el otro debe ser un mínimo. Sin embargo, si hay más de dos de dichos puntos, alguno puede ser punto silla. Además, si la curva C o la superficie S no está acotada, entonces f no necesariamente tiene máximos o mínimos.

Ejemplo. Halla el valor mínimo tomado por la función

$$f(x, y) = x^2 + (y - 2)^2 \quad \text{en la hipérbola } x^2 - y^2 = 1.$$

Solución:

El mínimo que se nos pide es sencillamente el cuadrado de la distancia del punto $(0, 2)$ a la hipérbola, luego está claro que existe. Siendo la curva no acotada, en este caso no existe máximo.

Para aplicar el teorema de Lagrange, teorema 11.9, hacemos

$$g(x, y) = x^2 - y^2 - 1, \quad \text{con } c = 0.$$

Entonces queremos maximizar y minimizar

$$f(x, y) = x^2 + (y - 2)^2 \quad \text{con la restricción } g(x, y) = x^2 - y^2 - 1 = 0.$$

Los gradientes son

$$\vec{\nabla} f(x, y) = 2x\vec{i} + 2(y - 2)\vec{j}, \quad \vec{\nabla} g(x, y) = 2x\vec{i} - 2y\vec{j}.$$

La condición (11.5), $\vec{\nabla} f(x_0) = \lambda \vec{\nabla} g(x_0)$ nos da dos ecuaciones, que junto con la ecuación que da la restricción, nos permite escribir un sistema de tres ecuaciones en las tres incógnitas x, y, λ :

$$\begin{cases} 2x = 2\lambda x \\ 2(y - 2) = -2\lambda y \\ x^2 - y^2 = 1 \end{cases} \Rightarrow \begin{cases} x = \lambda x \\ y - 2 = -\lambda y \\ x^2 - y^2 = 1 \end{cases} \stackrel{x \neq 0}{\Rightarrow} \begin{cases} \lambda = 1 \\ y = 1 \\ x = \pm\sqrt{2} \end{cases} \Rightarrow$$

$$\left(-\sqrt{2}, 1\right), \quad \left(\sqrt{2}, 1\right),$$

son los únicos puntos que pueden dar lugar a un valor mínimo. En cada uno de estos puntos la función f toma el valor 3, que es el valor mínimo que queríamos.

Nota. Podríamos haber resuelto este último problema de una manera más sencilla, escribiendo la restricción como $x^2 = 1 + y^2$ y eliminando la x en la función f por sustitución. Tendríamos que haber minimizado la función de una variable real $2y^2 - 4y + 5$.

11.3. EXTREMOS CON RESTRICCIONES Y MULTIPLICADORES DE LAGRANGE.

Observación. Si una superficie S está definida por un cierto número k de restricciones,

$$\begin{cases} g_1(x_1, \dots, x_n) = c_1 \\ g_2(x_1, \dots, x_n) = c_2 \\ \vdots \\ g_k(x_1, \dots, x_n) = c_k \end{cases}$$

entonces se puede generalizar el teorema de los multiplicadores de Lagrange:

Teorema 11.10 Si f tiene un máximo o un mínimo en $\vec{x}_0 \in S$, y si los vectores $\vec{\nabla} g_1(\vec{x}_0) \neq \vec{0}, \dots, \vec{\nabla} g_k(\vec{x}_0) \neq \vec{0}$, entonces deben existir constantes $\lambda_1, \dots, \lambda_k$ tales que

$$\vec{\nabla} f(\vec{x}_0) = \lambda_1 \vec{\nabla} g_1(\vec{x}_0) + \dots + \lambda_k \vec{\nabla} g_k(\vec{x}_0).$$

Damos un ejemplo de cómo se puede usar esta formulación general.

Ejemplo. Halla los puntos extremos de la función

$$f(x, y, z) = x + y + z \quad \text{sujetos a las restricciones} \quad x^2 + y^2 = 2 \quad \text{y} \quad x + z = 1.$$

Solución:

En este caso hay dos restricciones

$$g_1(x, y) = x^2 + y^2 - 2 = 0 \quad \text{y} \quad g_2(x, y) = x + z - 1 = 0.$$

Así, por el teorema generalizado de Lagrange, teorema 11.10, debemos hallar x, y, z, λ_1 y λ_2 tales que

$$\begin{cases} \vec{\nabla} f(x_0) = \lambda_1 \vec{\nabla} g_1(x_0) + \lambda_2 \vec{\nabla} g_2(x_0), \\ g_1(x, y) = x^2 + y^2 - 2 = 0, \\ g_2(x, y) = x + z - 1 = 0. \end{cases}$$

Calculando los gradientes

$$\vec{\nabla} f(x_0) = (1, 1, 1), \quad \vec{\nabla} g_1(x_0) = (2x, 2y, 0) \quad \text{y} \quad \vec{\nabla} g_2(x_0) = (1, 0, 1)$$

e igualando las componentes obtenemos un sistema de cinco ecuaciones

$$\begin{cases} 1 = \lambda_1 \cdot (2x) + \lambda_2 \cdot 1 \\ 1 = \lambda_1 \cdot (2y) + \lambda_2 \cdot 0 \\ 1 = \lambda_1 \cdot 0 + \lambda_2 \cdot 1 \\ x^2 + y^2 - 2 = 0 \\ x + z - 1 = 0 \end{cases} \Rightarrow \begin{cases} \lambda_2 = 1 \\ 2x \lambda_1 = 0 \\ 2y \lambda_1 = 1 \\ x^2 + y^2 - 2 = 0 \\ x + z - 1 = 0 \end{cases}$$

Como la tercera ecuación implica que $\lambda_1 \neq 0$, tendremos de la segunda $x = 0$. Entonces

$$\begin{cases} \lambda_2 = 1 \\ x = 0 \\ y = \pm\sqrt{2} \\ \lambda_1 = \frac{1}{2y} = \pm\frac{1}{2\sqrt{2}} \\ z = 1 \end{cases} \Rightarrow \begin{matrix} (0, \sqrt{2}, 1) & \text{es máximo,} \\ (0, -\sqrt{2}, 1) & \text{es mínimo.} \end{matrix}$$

Ejercicios.

1) Halla los extremos de f sujetos a las restricciones dadas.

- a) $f(x, y, z) = x - y + z$, si $x^2 + y^2 + z^2 = 2$;
- b) $f(x, y) = x$, si $x^2 + 2y^2 = 3$;
- c) $f(x, y) = x - y$, si $x^2 - y^2 = 2$;
- d) $f(x, y) = x^2 - y^2$, si $y = \cos x$;
- e) $f(x, y, z) = x^2 + y^2 + z^2$, si $z \geq 2 + x^2 + y^2$;
- f) $f(x, y) = x^2 + y^2$, en la hipérbola $xy = 1$;
- g) $f(x, y) = xy$, en la elipse $b^2 x^2 + a^2 y^2 = a^2 b^2$;
- h) $f(x, y) = xy^2$, en la elipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$;
- i) $f(x, y) = xy^2$, en la circunferencia $x^2 + y^2 = 1$;
- j) $f(x, y) = x^2 + y^2$, en la curva $x^4 + 7x^2 y^2 + y^4 = 1$;
- k) $f(x, y, z) = 2x + 3y + 5z$, en la esfera $x^2 + y^2 + z^2 = 19$;
- l) $f(x, y, z) = xyz$, en el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 19$.

2) Usa el método de los multiplicadores de Lagrange para hallar los valores mínimo y máximo absoluto en el disco unitario $D = \{(x, y) / x^2 + y^2 \leq 1\}$ de las funciones siguientes.

- a) $f(x, y) = x^2 + y^2 - x - y + 1$;
- b) $f(x, y) = x^2 + xy + y^2$.

11.4 Métodos numéricos de optimización con y sin restricciones.

En secciones anteriores, hemos estudiado el problema de buscar los extremos de funciones con valores reales. En el estudio de métodos numéricos este problema se conoce con el nombre de **problema de optimización**, y consiste en minimizar una función objetivo sujeta a algunas restricciones impuestas sobre las variables. Esto es, el problema general será del tipo **optimización no lineal con restricción**:

$$\begin{aligned}
 &\text{ONLCR} && \text{minimizar } F(\vec{x}) \quad \forall \vec{x} \in \mathbb{R}^n \\
 & && \text{sujeto a las restricciones} \\
 & && c_i = 0, i = 1, 2, \dots, m'; \\
 & && c_i \geq 0, i = m' + 1, \dots, m.
 \end{aligned}
 \tag{11.6}$$

11.4. MÉTODOS NUMÉRICOS DE OPTIMIZACIÓN CON Y SIN RESTRICCIONES.

Definición. Cualquier punto que satisface todas las restricciones del problema ONLCR (11.6) se dice que es *factible*.

Sea \bar{x}^* un punto factible para el problema ONLCR (11.6); definimos $N(\bar{x}^*, \delta)$ como el conjunto de todos los puntos factibles contenidos en un entorno de \bar{x}^* .

Definición. El punto \bar{x}^* es un **mínimo local fuerte** de ONLCR si existe un $\delta > 0$ tal que $F(\bar{x}^*) < F(\bar{y})$ para todo $\bar{y} \in N(\bar{x}^*, \delta)$, con $\bar{y} \neq \bar{x}^*$. El punto \bar{x}^* es un **mínimo local débil** de ONLCR si existe un $\delta > 0$ tal que $F(\bar{x}^*) \leq F(\bar{y})$ para todo $\bar{y} \in N(\bar{x}^*, \delta)$, con $\bar{y} \neq \bar{x}^*$.

Desafortunadamente, estas definiciones no son satisfactorias desde un punto de vista numérico: sería necesario evaluar $F(\bar{x}^*)$ en un conjunto (en general infinito) de puntos factibles. Sin embargo, si $F(\bar{x})$ y las restricciones $c_i(\bar{x})$ poseen algunas propiedades particulares, es posible dar otras condiciones más prácticas para caracterizar un punto de mínimo.

Consideremos primero el caso del problema de minimización **sin restricciones** en n dimensiones.

$$\text{SR} \quad \text{minimizar } F(\bar{x}) \quad \forall \bar{x} \in \mathbb{R}^n \quad (11.7)$$

Usando el desarrollo de Taylor de la función F alrededor de \bar{x}^* podemos escribir

$$F(\bar{x}^* + \varepsilon \bar{p}) = F(\bar{x}^*) + \varepsilon \bar{p}^t g(\bar{x}^*) + \frac{1}{2} \varepsilon^2 \bar{p}^t G(\bar{x}^* + \varepsilon \theta \bar{p}) \bar{p},$$

donde $0 \leq \theta \leq 1$, ε es un escalar y \bar{p} es un vector de \mathbb{R}^n .

Las condiciones necesarias para que el punto \bar{x}^* sea un mínimo local del problema SR (11.7) son

N.1 $\|g(\bar{x}^*)\| = 0$, esto es, \bar{x}^* es un punto estacionario;

N.2 $G(\bar{x}^*)$ es semidefinida positiva.

Por otro lado, las condiciones suficientes son

S.1 $\|g(\bar{x}^*)\| = 0$;

S.2 $G(\bar{x}^*)$ es definida positiva.

(Las demostraciones se pueden hacer por reducción al absurdo).

Un concepto muy importante en la construcción de un algoritmo es la medida del progreso; esto es, para un método iterativo es importante poseer un criterio razonable para decidir si la nueva aproximación es mejor que la anterior.

En el caso del problema de optimización sin restricciones, una medida natural del progreso viene dada por el valor de la función objetivo: parece razonable pedir una disminución de F a cada iteración, e imponer la **condición de descenso** $F(\bar{x}_{k+1}) < F(\bar{x}_k)$ para toda iteración $k \geq 0$. Todos los métodos de descenso son de la forma del siguiente algoritmo modelo:

Algoritmo modelo para el problema de minimización sin restricciones.

Sea \vec{x}_k la aproximación actual de \vec{x}^* :

U.1 Chequeo de la convergencia: si las condiciones para la convergencia se satisfacen, el algoritmo termina con \vec{x}_k como solución;

U.2 Cálculo de la dirección de búsqueda: se calcula un n -vector \vec{p}_k no nulo;

U.3 Cálculo de la longitud del paso: se calcula un escalar positivo α_k , la longitud del paso, para el cual se satisfaga la condición $F(\vec{x}_k + \alpha_k \vec{p}_k) < F(\vec{x}_k)$;

U.4 Renovación de la estimación del mínimo: se asignan los valores $\vec{x}_{k+1} \leftarrow \vec{x}_k + \alpha_k \vec{p}_k$ y $k \leftarrow k + 1$, y se vuelve al paso U1.

Es importante hacer notar que el paso U3 del algoritmo necesita de la resolución de un problema de una variable (encontrar el escalar α_k).

Para estar seguros de que se satisface la condición de descenso, \vec{p}_k y α_k tienen que satisfacer determinadas propiedades. Una manera estándar de garantizar que F se pueda reducir a la iteración k -ésima es pedir que \vec{p}_k sea una dirección descendente en \vec{x}_k , es decir, un vector que satisfaga

$$g^t(\vec{x}_k) \vec{p}_k < 0.$$

Si \vec{p}_k es una dirección descendente, entonces existirá $\alpha > 0$ tal que $F(\vec{x}_k + \alpha \vec{p}_k) < F(\vec{x}_k)$. Esta última condición por sí misma no es suficiente para asegurar que la sucesión $\{\vec{x}_k\}$ converja a un mínimo de F . Si $F(\vec{x}_k + \alpha \vec{p}_k) < F(\vec{x}_k)$ y la función F está acotada inferiormente, entonces claramente la sucesión $\{F(\vec{x}_k)\}$ converge, pero tendremos que asegurarnos de que converja a $F(\vec{x}^*)$.

Para asegurar la convergencia del algoritmo modelo se necesitan algunas condiciones sobre la función F . Antes de todo definamos el conjunto $L(\beta)$ para una función dada F y un escalar dado β : $L(\beta)$ es el conjunto de todos los puntos \vec{x} tales que $F(\vec{x}) \leq \beta$.

Si se satisfacen las siguientes condiciones:

- (i) F es C^2 , esto es, dos veces continuamente diferenciable,
- (ii) el conjunto $L(F(\vec{x}_0))$ es cerrado y acotado,
- (iii) F decrece "suficientemente" a cada iteración y
- (iv) el vector \vec{p}_k se mantiene no ortogonal al gradiente

entonces las iteraciones del algoritmo modelo satisfacen que

$$\lim_{k \rightarrow \infty} \|g(\vec{x}_k)\| = 0.$$

Este tipo de convergencia se denomina a veces *convergencia global* por el hecho de que no hay ninguna limitación sobre la cercanía de \vec{x}_0 al punto estacionario \vec{x}^* .

Remitimos al lector interesado en estos tipos de métodos a la literatura especializada.

Vamos ahora a considerar el problema de minimizar una función suave sujeta a un conjunto de restricciones del tipo igualdades lineales. Esto es, el problema

$$\begin{aligned} \text{IL} \quad & \text{minimizar } F(\vec{x}) \quad \forall \vec{x} \in \mathbb{R}^n \\ & \text{sujeto a las restricciones} \\ & \quad \hat{A} \vec{x} = \hat{b}, \end{aligned} \tag{11.8}$$

11.4. MÉTODOS NUMÉRICOS DE OPTIMIZACIÓN CON Y SIN RESTRICCIONES.

donde la i -ésima fila de la matriz \hat{A} , $m \times n$, contiene los coeficientes de la i -ésima restricción lineal

$$\hat{a}_{i1} x_1 + \cdots + \hat{a}_{in} x_n = \hat{b}_i.$$

Como para los problemas anteriores, el punto factible \bar{x}^* es un mínimo local de (11.8) sólo si $F(\bar{x}^*) \leq F(\bar{x})$ para todo punto factible \bar{x} en un entorno de \bar{x}^* . Está claro que no hay ningún punto factible si las restricciones son inconsistentes, por lo tanto asumimos que \hat{b} está en el rango de \hat{A} . Si t filas de la matriz \hat{A} son linealmente independientes, las restricciones eliminan t grados de libertad de la elección de \bar{x}^* . En dos dimensiones, por ejemplo, con la restricción $x_1 + x_2 = 0$, cualquier solución tendrá que estar en la recta $y = -x$. Es relevante notar que es importante el rango del conjunto de las restricciones, más que el número mismo de las restricciones. Volviendo al ejemplo anterior, si se añade la restricción $2x_1 + 2x_2 = 0$ (que es linealmente dependiente de la primera), no se puede eliminar ningún grado ulterior de libertad; al contrario, con la restricción $x_1 - x_2 = 1$ (que es linealmente independiente de la primera), las restricciones eliminan dos grados de libertad, y en este caso se determina la solución completamente (será la intersección de las dos rectas $y = -x$ e $y = x - 1$).

Dado que las restricciones constituyen un sistema de ecuaciones lineales, las propiedades de los subespacios lineales hacen posible una caracterización de todos los "desplazamientos factibles" posibles. Si \bar{x} y \hat{x} son dos puntos factibles, entonces $\hat{A}(\bar{x} - \hat{x}) = 0$, esto es, el paso \bar{p} de un punto factible a otro tiene que ser ortogonal a las filas de \hat{A} :

$$\hat{A}\bar{p} = 0. \quad (11.9)$$

Cualquier paso desde un punto factible a otro a lo largo de estas direcciones no incumple la restricción, dado que $\hat{A}(\hat{x} + \alpha\bar{p}) = \hat{A}\hat{x} = \hat{b}$.

La condición (11.9) define un subespacio de vectores cuyos vectores de base forman las columnas de una matriz Z con la característica que $\hat{A}Z = 0$ y cada dirección factible puede escribirse como combinación lineal de las columnas de Z . Luego si \bar{p} satisface (11.9), puede escribirse como

$$\bar{p} = Z\bar{p}_z \quad (11.10)$$

para algún vector \bar{p}_z .

Usando el desarrollo de Taylor de la función F alrededor de \bar{x}^* a lo largo de una dirección factible $\bar{p} = Z\bar{p}_z$, podemos escribir, con $0 \leq \theta \leq 1$ y ε un escalar positivo

$$F(\bar{x}^* + \varepsilon\bar{p}) = F(\bar{x}^* + \varepsilon Z\bar{p}_z) = F(\bar{x}^*) + \varepsilon\bar{p}_z^t Z^t g(\bar{x}^*) + \frac{1}{2}\varepsilon^2\bar{p}_z^t Z^t G(\bar{x}^* + \varepsilon\theta\bar{p}) Z\bar{p}_z.$$

Las condiciones necesarias para que el punto \bar{x}^* sea un mínimo local del problema IL (11.8) son

- N.1 $\hat{A}\bar{x}^* = \hat{b}$;
- N.2 $Z^t g(\bar{x}^*) = 0$;
- N.3 $Z^t G(\bar{x}^*) Z$ es semidefinida positiva.

Por otro lado, las condiciones suficientes son

- S.1 $\hat{A} \bar{x}^* = \hat{b}$;
- S.2 $Z^t g(\bar{x}^*) = 0$;
- S.3 $Z^t G(\bar{x}^*) Z$ es definida positiva.

Como ya se ha comentado, la imposición de t restricciones del tipo de igualdades lineales y que sean linealmente independientes, reduce la dimensión del problema de optimización a $n - t$. Usando la condición (11.10), el algoritmo modelo dado para el problema de minimización sin restricciones, se puede modificar para resolver el problema IL generando una sucesión de iteraciones factibles.

Algoritmo modelo para el problema de minimización con restricciones del tipo igualdades lineales.

Sea \bar{x}_k la aproximación actual de \bar{x}^* :

IL.1 Chequeo de la convergencia: si las condiciones para la convergencia se satisfacen, el algoritmo termina con \bar{x}_k como solución;

IL.2 Cálculo de la dirección de búsqueda factible: se calcula un $(n - t)$ -vector \vec{p}_k no nulo dado por (11.10)

$$\vec{p} = Z \vec{p}_z; \tag{11.10}$$

IL.3 Cálculo de la longitud del paso: se calcula un escalar positivo α_k , la longitud del paso, para el cual se satisfaga la condición $F(\bar{x}_k + \alpha_k \vec{p}_k) < F(\bar{x}_k)$;

IL.4 Renovación de la estimación del mínimo: se asignan los valores $\bar{x}_{k+1} \leftarrow \bar{x}_k + \alpha_k \vec{p}_k$ y $k \leftarrow k + 1$, y se vuelve al paso IL1.

Como el lector ya imagina, en la literatura especializada se estudian problemas más complicados como minimizar una función suave sujeta a un conjunto de restricciones del tipo **desigualdades lineales**. Esto es, el problema

$$\begin{aligned} \text{DL} \quad & \text{minimizar } F(\bar{x}) \quad \forall \bar{x} \in \mathbb{R}^n \\ & \text{sujeto a las restricciones} \\ & \hat{A} \bar{x} \geq \hat{b}. \end{aligned} \tag{11.11}$$

y también los problemas mencionados al principio de la sección de restricciones no lineales, en particular, el problema de minimizar una función suave sujeta a un conjunto de restricciones del tipo **igualdades no lineales**:

$$\begin{aligned} \text{INL} \quad & \text{minimizar } F(\bar{x}) \quad \forall \bar{x} \in \mathbb{R}^n \\ & \text{sujeto a las restricciones} \\ & \hat{c}(\bar{x}) = 0, \quad i = 1, \dots, t; \end{aligned} \tag{11.12}$$

y el problema de minimizar una función suave sujeta a un conjunto de restricciones del tipo **desigualdades no lineales**:

$$\begin{aligned} \text{DNL} \quad & \text{minimizar } F(\bar{x}) \quad \forall \bar{x} \in \mathbb{R}^n \\ & \text{sujeto a las restricciones} \\ & \hat{c}(\bar{x}) \geq 0, \quad i = 1, \dots, m. \end{aligned} \tag{11.13}$$

Como antes, remitimos al lector interesado en estos tipos de métodos a la literatura especializada.

Capítulo 12

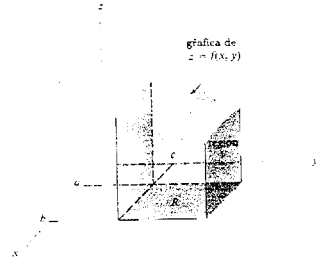
Integración.

12.1 Introducción.

En capítulos previos hemos estudiado las integrales ordinarias $\int_a^b f(x) dx$. Ahora vamos a estudiar las integrales dobles $\iint_{\Omega} f(x, y) dx dy$ y un posteriormente las integrales triples $\iiint_V f(x, y, z) dx dy dz$.

Sea $f: R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una función continua de dos variables, cuyo dominio R es un rectángulo con lados paralelos a los ejes coordenados, que puede escribirse en términos de dos intervalos cerrados $[a, b]$ y $[c, d]$ como producto cartesiano $R = [a, b] \times [c, d]$.

Supongáse que $f(x, y) \geq 0$ en R , de manera que la gráfica de $z = f(x, y)$ es una superficie que está encima del rectángulo R . Esta superficie, el rectángulo R y los cuatro planos $x = a$, $x = b$, $y = c$ e $y = d$ forman la frontera de una región V en el espacio. En la sección 12.2 definiremos de manera rigurosa el volumen V por medio del método de las sumas de Riemann. Sin embargo, para tener un conocimiento intuitivo de este método supongamos que se ha definido el volumen de una región.



Definición. El volumen de la región por encima de R y debajo de la gráfica de f se llama la **integral doble de f sobre R** y se denota por

$$\int_R f, \quad \int_R f(x, y) dA, \quad \int_R f(x, y) dx dy, \quad \iint_R f(x, y) dx dy.$$

Ejemplo 1. Si f está definida por $f(x, y) = k$, donde k es una constante positiva, entonces

$$\int_R f(x, y) dx dy = k(b - a)(d - c),$$

pues la integral es igual al volumen de un paralelepípedo con base R y altura k .

Ejemplo 2. Si $f(x, y) = 1 - x$ y $R = [0, 1] \times [0, 1]$, entonces

$$\int_R f(x, y) dx dy = \frac{1}{2},$$

pues la integral es igual al volumen del sólido triangular obtenido dividiendo un cubo con un plano que pasa por la diagonal.

Estas ideas son similares a la de la integral simple $\int_a^b f(x) dx$, que representa al área

bajo la gráfica de f si $f \geq 0$ y continua. Podemos aproximar $\int_a^b f(x) dx$ escogiendo una partición P de $[a, b]$ tales que $a = x_0 < x_1 < x_2 < \dots < x_n = b$, seleccionando puntos $c_i \in [x_i, x_{i+1}]$ y formando la suma de Riemann (ver la sección 3.1)

$$\sum_{i=0}^{n-1} f(c_i) (x_{i+1} - x_i) \approx \int_a^b f(x) dx.$$

Hay un método para calcular volúmenes conocido como **principio de Cavalieri**.

Supongamos que tenemos un cuerpo sólido y denotemos por $A(x)$ el área de su sección transversal medida a una distancia x de un plano de referencia. De acuerdo con el principio de Cavalieri, el volumen del cuerpo está dado por

$$\text{Volumen} = \int_a^b A(x) dx,$$

donde a y b son las distancias mínima y máxima a partir del plano de referencia.

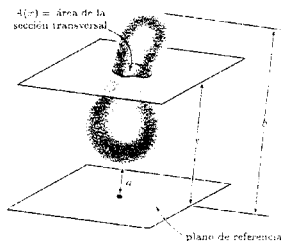
Si consideramos una partición P de $[a, b]$ tal que $a = x_0 < x_1 < x_2 < \dots < x_n = b$, entonces una suma de Riemann que se aproxima al valor de la integral es

$$\text{Volumen} = \int_a^b A(x) dx \approx \sum_{i=0}^{n-1} A(c_i) (x_{i+1} - x_i).$$

Pero esta suma también se aproxima al volumen del cuerpo, pues $A(x) \Delta x$ es el volumen de una sección con área transversal $A(x)$ y ancho Δx .

Ahora usamos el principio de Cavalieri para evaluar integrales dobles. Consideremos la región sólida bajo la gráfica $z = f(x, y)$ definida en la región $[a, b] \times [c, d]$, donde f es continua y mayor que cero. Hay dos funciones naturales para el área de sección transversal:

- a) usando planos perpendiculares al eje x ,
- b) usando planos perpendiculares al eje y .



12.1. INTRODUCCIÓN.

La sección transversal determinada por un plano $x = x_0$, del primer tipo, es la región plana debajo de la gráfica $z = f(x, y)$ de $y = c$ a $y = d$. Cuando fijamos $x = x_0$, tenemos la función $y \rightsquigarrow f(x_0, y)$ que es continua en $[c, d]$. Entonces, el área de la sección transversal es $A(x_0) = \int_c^d f(x_0, y) dy$. Así, la función A de área de sección transversal tiene dominio $[a, b]$ y

$$A: x \rightsquigarrow \int_c^d f(x, y) dy.$$

Por el principio de Cavalieri, el volumen V de la región debajo de $z = f(x, y)$ debe ser igual a

$$V = \int_R f(x, y) dA = \int_a^b A(x) dx = \int_a^b \left[\int_c^d f(x, y) dy \right] dx. \quad (12.1)$$

La última integral de la fórmula (12.1), $\int_a^b \left[\int_c^d f(x, y) dy \right] dx$ se conoce como **integral iterada**, pues se obtiene integrando respecto a la variable y y después integrando el resultado respecto a la variable x .

Si invertimos los papeles de x e y , y consideramos planos perpendiculares al eje y del segundo tipo, obtenemos

$$V = \int_R f(x, y) dA = \int_c^d A(y) dy = \int_c^d \left[\int_a^b f(x, y) dx \right] dy. \quad (12.2)$$

La expresión a la derecha de la fórmula (12.2), $\int_c^d \left[\int_a^b f(x, y) dx \right] dy$ es la integral iterada obtenida integrando primero respecto a la variable x y después integrando el resultado respecto a la variable y .

Como veremos más adelante en el teorema de Fubini 12.3, las dos fórmulas (12.1) y (12.2) deberán ser válidas al mismo tiempo. Como veremos en los ejemplos siguientes, las integrales iteradas de las ecuaciones (12.1) y (12.2) proporcionan un método poderoso para calcular la integral doble de una función de dos variables.

Ejemplo 1. Sea $z = f(x, y) = x^2 + y^2$ y sea $R = [-1, 1] \times [0, 1]$. Evalúa la integral

$$\iint_R (x^2 + y^2) dx dy.$$

Solución:

Por la ecuación (12.2) tenemos

$$\iint_R (x^2 + y^2) dx dy = \int_0^1 \left[\int_{-1}^1 (x^2 + y^2) dx \right] dy.$$

Para evaluar $\int_{-1}^1 (x^2 + y^2) dx$, tratamos y como constante e integramos respecto a x .

Como $F(x, y) = \frac{x^3}{3} + y^2 x$ es una primitiva de $f(x, y) = x^2 + y^2$ (aquí y es considerada constante) podemos integrar usando métodos de cálculo de una variable, y obtener

$$\int_{-1}^1 (x^2 + y^2) dx = \left[\frac{x^3}{3} + y^2 x \right]_{-1}^1 = \frac{2}{3} + 2y^2.$$

Después integramos esta nueva función con respecto a y , de 0 a 1, para obtener

$$\int_0^1 \left(\frac{2}{3} + 2y^2 \right) dy = \left[\frac{2}{3}y + \frac{2}{3}y^3 \right]_0^1 = \frac{4}{3}.$$

Para completar, ahora evaluamos la integral $\iint_R (x^2 + y^2) dx dy$ usando la fórmula (12.1)

$$\iint_R (x^2 + y^2) dx dy = \int_{-1}^1 \left[\int_0^1 (x^2 + y^2) dy \right] dx.$$

Para evaluar $\int_0^1 (x^2 + y^2) dy$, tratamos x como constante e integramos respecto a y .

Como $F(x, y) = x^2 y + \frac{y^3}{3}$ es una primitiva de $f(x, y) = x^2 + y^2$ (aquí x se considera constante) podemos integrar usando métodos de cálculo de una variable, y obtener

$$\int_0^1 (x^2 + y^2) dy = \left[x^2 y + \frac{y^3}{3} \right]_0^1 = x^2 + \frac{1}{3}.$$

Después integramos esta nueva función con respecto a x , de -1 a 1 , para obtener

$$\int_{-1}^1 \left(x^2 + \frac{1}{3} \right) dx = \left[\frac{x^3}{3} + \frac{x}{3} \right]_{-1}^1 = \frac{4}{3}.$$

Ejemplo 2. Calcula $\int_S \cos x \operatorname{sen} y dx dy$, donde S es el cuadrado $\left[0, \frac{\pi}{2}\right] \times \left[0, \frac{\pi}{2}\right]$.

Solución:

Por la fórmula (12.2) tenemos

$$\begin{aligned} \int_S \cos x \operatorname{sen} y dx dy &= \int_0^{\pi/2} \left[\int_0^{\pi/2} \cos x \operatorname{sen} y dx \right] dy = \\ &= \int_0^{\pi/2} \operatorname{sen} y \left[\int_0^{\pi/2} \cos x dx \right] dy = \int_0^{\pi/2} \operatorname{sen} y [\operatorname{sen} x]_0^{\pi/2} dy = \\ &= \int_0^{\pi/2} \operatorname{sen} y dy = [-\cos x]_0^{\pi/2} = 1. \end{aligned}$$

Compruebe el lector que se obtiene el mismo resultado usando la fórmula (12.1).

12.2. LA INTEGRAL DOBLE SOBRE UN RECTÁNGULO.

Ejercicios.

1) Evalúa las siguientes integrales iteradas y luego cambia el orden de integración

$$a) \int_{-1}^1 \int_0^1 (x^4 y + y^2) dy dx;$$

$$b) \int_0^{\pi/2} \int_0^1 (y \cos x + 2) dy dx;$$

$$c) \int_0^1 \int_0^1 (x y e^{x+y}) dy dx;$$

$$d) \int_{-1}^1 \int_1^2 (-x \log y) dy dx.$$

2) Evalúa las integrales dobles siguientes, donde R es el rectángulo $[0, 2] \times [-1, 0]$.

$$a) \int_R \left[|y| \cos \left(\frac{\pi x}{4} \right) \right] dy dx;$$

$$b) \int_R \left[-x e^x \sin \left(\frac{\pi y}{2} \right) \right] dy dx;$$

$$c) \int_R (x^2 y^2 + x) dy dx;$$

$$d) \int_R (x^2 + y^2) dy dx.$$

3) Halla el volumen acotado por la gráfica de la función $f(x, y)$ en el rectángulo R indicados y los cuatro planos verticales del rectángulo R .

$$a) f(x, y) = 1 + 2x + 3y, \quad R = [1, 2] \times [0, 1];$$

$$b) f(x, y) = x^4 + y^2, \quad R = [-1, 1] \times [-3, -2].$$

12.2 La integral doble sobre un rectángulo.

Estamos ahora preparados para dar una definición rigurosa de la integral doble como límite de una sucesión de sumas. No requeriremos que $f(x, y) \geq 0$; pero si $f(x, y)$ toma valores negativos, interpretaremos la integral como un volumen con signo, como para el área bajo la gráfica de una función de una variable.

Consideremos un rectángulo $R \subset \mathbb{R}^2$, que es el producto cartesiano $R = [a, b] \times [c, d]$.

Definición. Por una **partición regular** P de R de orden n , entendemos dos colecciones ordenadas de $n + 1$ puntos igualmente espaciados $\{x_j\}_{j=0}^n$ y $\{y_k\}_{k=0}^n$; esto es, puntos que satisfacen

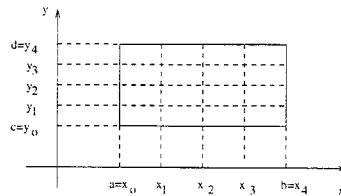
$$a = x_0 < x_1 < x_2 < \dots < x_n = b,$$

$$c = y_0 < y_1 < y_2 < \dots < y_n = d$$

y

$$\Delta x = x_{j+1} - x_j = \frac{b - a}{n},$$

$$\Delta y = y_{k+1} - y_k = \frac{d - c}{n}.$$



Sea R_{jk} el rectángulo $[x_j, x_{j+1}] \times [y_k, y_{k+1}]$, y sea \bar{c}_{jk} cualquier punto en R_{jk} . Supongamos que $f: R \rightarrow \mathbb{R}$ es una función acotada con valores reales. Podemos entonces formar la suma

$$S_n = \sum_{j,k=0}^{n-1} f(\bar{c}_{jk}) \Delta x \Delta y = \sum_{j,k=0}^{n-1} f(\bar{c}_{jk}) \Delta A,$$

donde $\Delta A = \Delta x \Delta y$.

Esta suma está realizada sobre todo índice j y k de 0 a $n - 1$, de modo que hay n^2 términos. Una suma de este tipo se llama suma de Riemann para f .

Definición. Si la sucesión $\{S_n\}$ converge a un límite S cuando $n \rightarrow \infty$ y el límite S es el mismo para cualquier selección de puntos \bar{c}_{jk} en los rectángulos R_{jk} , entonces decimos que la función f es **integrable sobre R** y escribimos

$$\int_R f, \quad \int_R f(x, y) dA, \quad \int_R f(x, y) dx dy, \quad \iint_R f(x, y) dx dy$$

para el límite S ; es decir

$$\iint_R f(x, y) dx dy = \lim_{n \rightarrow \infty} S_n = S = \lim_{n \rightarrow \infty} \sum_{j,k=0}^{n-1} f(\bar{c}_{jk}) \Delta x \Delta y.$$

En la sección 3.1 se escogieron los valores mínimo y máximo de la función $f(x)$ para obtener el área del recinto por defecto y por exceso, respectivamente. De manera similar podemos considerar una partición genérica P del rectángulo R considerando una partición de $[a, b]$, $P_1 = \{x_0, x_1, \dots, x_m\}$ y una partición de $[c, d]$, $P_2 = \{y_0, y_1, \dots, y_n\}$. La partición $P = P_1 \times P_2$ divide el rectángulo R en $m \times n$ rectángulos que no se solapan, $R_{jk} = \{(x, y) / x_{j-1} < x < x_j, y_{k-1} < y < y_k\}$. En cada rectángulo R_{jk} la función $f(x, y)$ toma el valor máximo M_{jk} y el valor mínimo m_{jk} . Entonces, podemos definir la **suma superior de f** relativa a la partición P como

$$U_f(P) = \sum_{j,k=0}^{n-1} M_{jk} \Delta x_j \Delta y_k,$$

y la **suma inferior de f** relativa a la partición P como

$$L_f(P) = \sum_{j,k=0}^{n-1} m_{jk} \Delta x_j \Delta y_k.$$

Si $f(x, y) \geq 0$, la existencia del $\lim_{n \rightarrow \infty} S_n$ tiene un significado geométrico directo. Consideremos la gráfica de $z = f(x, y)$ como la tapa de un sólido cuya base es el rectángulo R . La suma superior de f , $U_f(P) = \sum_{j,k=0}^{n-1} M_{jk} \Delta x_j \Delta y_k$, representa el volumen del sólido

12.2. LA INTEGRAL DOBLE SOBRE UN RECTÁNGULO.

circunscrito; mientras que la suma inferior de f , $L_f(P) = \sum_{j,k=0}^{n-1} m_{jk} \Delta x_j \Delta y_k$, es igual al volumen de un sólido inscrito. Será entonces

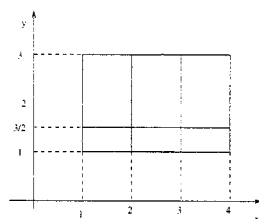
$$L_f(P) \leq S \leq U_f(P).$$

Por lo tanto, si existe el límite $\lim_{n \rightarrow \infty} S_n$ y es independiente de los puntos de R_{jk} , se sigue que los volúmenes de los sólidos inscrito y circunscrito tienden al mismo límite cuando $n \rightarrow \infty$.

Ejemplo. Consideremos en el rectángulo $R = \{(x, y) / 1 \leq x \leq 4, 1 \leq y \leq 3\}$ la función

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

Como partición de $[1, 4]$ tomamos $P_1 = \{1, 2, 3, 4\}$ y como partición de $[1, 3]$ tomamos $P_2 = \{1, \frac{3}{2}, 3\}$. La partición $P = P_1 \times P_2$ divide el rectángulo inicial en seis rectángulos. En cada uno de los rectángulos R_{jk} , la función alcanza su máximo M_{jk} en el punto (x_j, y_k) , el vértice más alejado del origen,



$$M_{jk} = f(x_j, y_k) = x_j + y_k - 2;$$

y así mismo la función alcanza su mínimo m_{jk} en el punto (x_{j-1}, y_{k-1}) , el vértice más próximo al origen.

$$m_{jk} = f(x_{j-1}, y_{k-1}) = x_{j-1} + y_{k-1} - 2.$$

Luego

$$\begin{aligned} U_f(P) &= \sum_{j,k=0}^{n-1} M_{jk} \Delta x_j \Delta y_k = \\ &= \frac{3}{2} \left(\frac{1}{2}\right) + 3 \left(\frac{3}{2}\right) + \frac{5}{2} \left(\frac{1}{2}\right) + 4 \left(\frac{3}{2}\right) + \frac{7}{2} \left(\frac{1}{2}\right) + 5 \left(\frac{3}{2}\right) = \frac{87}{4}, \end{aligned}$$

$$\begin{aligned} L_f(P) &= \sum_{j,k=0}^{n-1} m_{jk} \Delta x_j \Delta y_k = \\ &= 0 \left(\frac{1}{2}\right) + \frac{1}{2} \left(\frac{3}{2}\right) + 1 \left(\frac{1}{2}\right) + \frac{3}{2} \left(\frac{3}{2}\right) + 2 \left(\frac{1}{2}\right) + \frac{5}{2} \left(\frac{3}{2}\right) = \frac{33}{4}. \end{aligned}$$

Si usamos las fórmulas de las integrales iteradas (12.1) y (12.2) obtenemos

$$\begin{aligned} \iint_R (x + y - 2) dA &= \int_1^4 \int_1^3 (x + y - 2) dy dx = \int_1^4 \left[\int_1^3 (x + y - 2) dy \right] dx = \\ &= \int_1^4 \left[xy + \frac{y^2}{2} - 2y \right]_1^3 dx = \int_1^4 (2x) dx = [x^2]_1^4 = 15; \end{aligned}$$

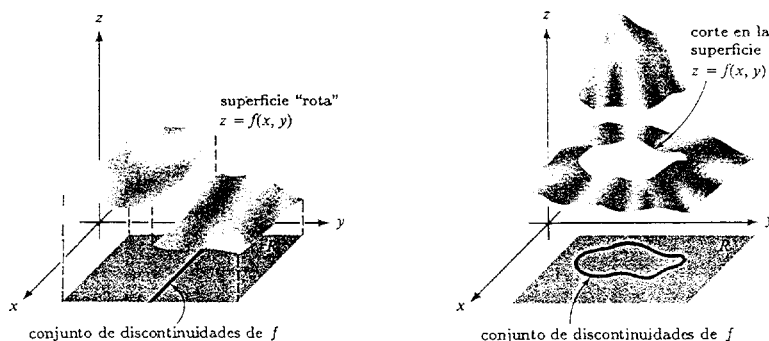
$$\begin{aligned} \iint_R (x+y-2) dA &= \int_1^3 \int_1^4 (x+y-2) dx dy = \int_1^3 \left[\int_1^4 (x+y-2) dx \right] dy = \\ &= \int_1^3 \left[\frac{x^2}{2} + xy - 2x \right]_1^4 dy = \int_1^3 \left(3y + \frac{3}{2} \right) dy = \left[\frac{3}{2}y^2 + \frac{3}{2}y \right]_1^3 = 15. \end{aligned}$$

Como se esperaba, $\frac{33}{4} \leq 15 \leq \frac{87}{4}$.

Hay un teorema que no demostraremos que garantiza la existencia de la integral de toda función continua.

Teorema 12.1 *Cualquier función continua definida en un rectángulo R es integrable.*

También hay otro teorema que garantiza la existencia de la integral de ciertas funciones discontinuas. Nos interesaremos en especial por funciones cuyas discontinuidades estén en curvas en el plano xy . Es decir, f es continua en cada punto que esté en R pero no sobre la curva. Curvas útiles son las gráficas de funciones $y = \phi(x)$, con $a \leq x \leq b$, o $x = \psi(y)$ con $c \leq y \leq d$, o unión finita de dichas gráficas.



Teorema 12.2 *Sea $f: R \rightarrow \mathbb{R}$ una función acotada con valores reales, definida en el rectángulo R . Supóngase que el conjunto de puntos donde f es discontinua esté formado por una unión finita de gráficas de funciones continuas. Entonces f es integrable sobre el rectángulo R .*

Nota. Recordar que una función está acotada si existe un número $M > 0$ tal que $-M \leq f(x, y) \leq M$ para todo (x, y) en el dominio de f . Una función continua en un rectángulo cerrado siempre está acotada, pero, por ejemplo, $f(x, y) = x^{-1}$ en $(0, 1] \times [0, 1]$ no está acotada, pues x^{-1} se vuelve arbitrariamente grande para valores x cerca de cero.

Geoméricamente, el teorema 12.2 implica que si una función no negativa f “no se porta demasiado mal”, entonces los volúmenes de los sólidos circunscrito e inscrito aproximarán al “verdadero” volumen bajo su gráfica.

12.2. LA INTEGRAL DOBLE SOBRE UN RECTÁNGULO.

De la definición de integral como límite de sumas y los teoremas de límites, se pueden deducir algunas propiedades fundamentales de la integral doble, que son esencialmente las mismas que para la integral de una función con valores reales, de una variable.

Propiedades de la integral doble.

Las siguientes propiedades elementales de la integral doble son análogas a las que vimos en el caso de una variable. Se supone que las funciones f y g son continuas en R .

1) Linealidad.

$$\int_R [f(x, y) + g(x, y)] dA = \int_R f(x, y) dA + \int_R g(x, y) dA;$$

2) Homogeneidad.

$$\int_R k f(x, y) dA = k \int_R f(x, y) dA, \quad \forall k \in \mathbb{R};$$

3) Monotonía. Si $f(x, y) \geq g(x, y)$ en R , entonces

$$\int_R f(x, y) dA \geq \int_R g(x, y) dA;$$

4) Aditividad. Si R_i , $i = 1, \dots, m$ son rectángulos disjuntos entre sí, tales que f es integrable sobre cada R_i y si $Q = R_1 \cup R_2 \cup \dots \cup R_m$ es un rectángulo, entonces una función acotada $f: Q \rightarrow \mathbb{R}$ es integrable sobre Q y

$$\int_Q f(x, y) dA = \sum_{i=1}^m \int_{R_i} f(x, y) dA.$$

Demostración:

Las propiedades 1) y 2) son consecuencia de la definición como límite de una suma y los siguientes hechos acerca de series convergentes $\{S_n\}$ y $\{T_n\}$, que hemos visto en el teorema 7.2

$$\lim_{n \rightarrow \infty} (S_n + T_n) = \lim_{n \rightarrow \infty} S_n + \lim_{n \rightarrow \infty} T_n,$$

$$\lim_{n \rightarrow \infty} (cS_n) = c \lim_{n \rightarrow \infty} S_n.$$

Para demostrar la propiedad 3) de monotonía, observemos primero que si $h(x, y) \geq 0$ y $\{S_n\}$ es una sucesión de sumas de Riemann que converge a $\int_R h(x, y) dA$, entonces, $S_n \geq 0$ para todo n , de modo que la integral será también positiva o nula. Si $f(x, y) \geq g(x, y)$ para todo $(x, y) \in R$, entonces $(f - g)(x, y) \geq 0$ para todo (x, y) y usando las propiedades 1) y 2), tenemos

$$\int_R f(x, y) dA - \int_R g(x, y) dA = \int_R [f(x, y) - g(x, y)] dA \geq 0,$$

que prueba la propiedad 3). La demostración de la propiedad 4) es más técnica y no la presentamos, aunque es intuitivamente obvia.

c.q.d.

Otro resultado importante es la desigualdad

$$\left| \int_R f(x, y) dA \right| \leq \int_R |f(x, y)| dA.$$

Por la definición de valor absoluto, $-|f(x, y)| \leq f(x, y) \leq |f(x, y)|$; por lo tanto, de la monotonía y la homogeneidad de la integración (con $c = -1$),

$$-\int_R |f(x, y)| dA \leq \int_R f(x, y) dA \leq \int_R |f(x, y)| dA,$$

que es la fórmula anterior.

Aunque hemos visto la integrabilidad de varias funciones, no hemos establecido aún un método general para calcular integrales dobles. En el caso de funciones de una variable, evitábamos calcular la integral $\int_a^b f(x) dx$ como límite de sumas de Riemann, mediante el uso del teorema fundamental del cálculo integral, teorema 3.2, o regla de Barrow (3.1)

$$\int_a^b f(t) dt = F(x) \Big|_a^b = F(b) - F(a),$$

donde F es una primitiva de f , esto es $F' = f$.

Para funciones de dos variables esto no funciona; sin embargo, como hemos visto en la sección 12.1, a menudo es posible reducir una integral doble sobre un rectángulo, a integrales simples iteradas, a las cuales se aplica el teorema fundamental. El teorema de Fubini, que ya se mencionó en la sección anterior, justifica rigurosamente esta reducción a integrales iteradas.

Teorema 12.3 Teorema de Fubini.

Sea f una función continua con dominio rectangular $R = [a, b] \times [c, d]$. Entonces

$$\int_R f(x, y) dA = \int_a^b \int_c^d f(x, y) dy dx = \int_c^d \int_a^b f(x, y) dx dy. \quad (12.3)$$

Demostración:

Primero demostraremos que

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

Sea $c = y_0 < y_1 < \dots < y_n = d$ una partición de $[c, d]$ en n partes iguales. Definimos la función

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

12.2. LA INTEGRAL DOBLE SOBRE UN RECTÁNGULO.

Entonces, será

$$F(x) = \sum_{k=0}^{n-1} \int_{y_k}^{y_{k+1}} f(x, y) dy.$$

Usando el teorema del valor medio integral, teorema 3.3, para cada x fija y cada k tenemos

$$\int_{y_k}^{y_{k+1}} f(x, y) dy = f(x, Y_k(x)) (y_{k+1} - y_k),$$

donde el punto $Y_k(x)$ pertenece a $[y_k, y_{k+1}]$ y puede depender de x , k y n . Entonces, se tiene que

$$F(x) = \sum_{k=0}^{n-1} \int_{y_k}^{y_{k+1}} f(x, y) dy = \sum_{k=0}^{n-1} f(x, Y_k(x)) (y_{k+1} - y_k).$$

Ahora bien, por la definición de la integral en una variable como límite de sumas de Riemann,

$$\int_a^b F(x) dx = \int_a^b \left[\int_c^d f(x, y) dy \right] dx = \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} F(p_j) (x_{j+1} - x_j),$$

donde $a = x_0 < x_1 < \dots < x_n = b$ una partición de $[a, b]$ en n partes iguales y p_j es cualquier punto en $[x_j, x_{j+1}]$. Haciendo $\bar{c}_{jk} = (p_j, Y_k(p_j)) \in R_{jk}$, tenemos

$$F(p_j) = \sum_{k=0}^{n-1} f(p_j, Y_k(p_j)) (y_{k+1} - y_k) = \sum_{k=0}^{n-1} f(\bar{c}_{jk}) (y_{k+1} - y_k)$$

y por lo tanto

$$\begin{aligned} \int_a^b \int_c^d f(x, y) dy dx &= \int_a^b F(x) dx = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} F(p_j) (x_{j+1} - x_j) = \\ &= \lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(\bar{c}_{jk}) (y_{k+1} - y_k) (x_{j+1} - x_j) = \\ &= \int_R f(x, y) dA. \end{aligned}$$

Con el mismo razonamiento podemos demostrar que

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

c.q.d

El teorema de Fubini se puede generalizar al caso en que f no necesariamente sea continua. Aunque sin demostración, enunciamos la versión más general.

Teorema 12.4 Teorema de Fubini.

Sea f una función acotada con dominio rectangular $R = [a, b] \times [c, d]$, cuyas discontinuidades forman una unión finita de gráficas de funciones continuas. Si

$$\int_c^d f(x, y) dy \quad \text{existe para cada } x \in [a, b],$$

entonces existe la integral iterada $\int_a^b \left[\int_c^d f(x, y) dy \right] dx$, y además

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

De manera análoga, si

$$\int_a^b f(x, y) dx \quad \text{existe para cada } y \in [c, d],$$

entonces existe la integral iterada $\int_c^d \left[\int_a^b f(x, y) dx \right] dy$, y además

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

Si todas estas condiciones se cumplen simultáneamente, obtenemos la fórmula (12.3)

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

Ejemplo. Calcula $\int_R (x^2 + y) dx dy$, donde $R = [0, 1] \times [0, 1]$.

Solución:

Por el teorema de Fubini 12.3, por un lado se tiene que

$$\int_R (x^2 + y) dx dy = \int_0^1 \int_0^1 (x^2 + y) dx dy = \int_0^1 \left[\int_0^1 (x^2 + y) dx \right] dy.$$

Por el teorema fundamental del cálculo, teorema 3.2, se puede ejecutar la integración en x manteniendo la variable y fija:

$$\int_0^1 (x^2 + y) dx = \left[\frac{x^3}{3} + yx \right]_0^1 = \frac{1}{3} + y.$$

Ahora, integramos la función restante (sólo de y) respecto a y :

$$\int_R (x^2 + y) dx dy = \int_0^1 \left[\frac{1}{3} + y \right] dy = \left[\frac{1}{3}y + \frac{y^2}{2} \right]_0^1 = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}.$$

12.2. LA INTEGRAL DOBLE SOBRE UN RECTÁNGULO.

El teorema de Fubini dice que al intercambiar el orden de integración en las integrales iteradas, no se altera el resultado; de hecho

$$\int_R (x^2 + y) \, dx \, dy = \int_0^1 \int_0^1 (x^2 + y) \, dy \, dx = \int_0^1 \left[\int_0^1 (x^2 + y) \, dy \right] dx.$$

Nuevamente por el teorema fundamental del cálculo, se puede ejecutar la integración en y manteniendo la variable x fija:

$$\int_0^1 (x^2 + y) \, dy = \left[x^2 y + \frac{y^2}{2} \right]_0^1 = x^2 + \frac{1}{2}.$$

Ahora, integramos la función restante (sólo de x) respecto a x :

$$\int_R (x^2 + y) \, dx \, dy = \int_0^1 \left[x^2 + \frac{1}{2} \right] dx = \left[\frac{x^3}{3} + \frac{1}{2} x \right]_0^1 = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}.$$

Vimos que cuando $f(x, y) \geq 0$ en $R = [a, b] \times [c, d]$, la integral doble $\int_R f(x, y) \, dx \, dy$ se puede interpretar como un volumen. Si la función toma también valores negativos, entonces la integral doble se puede interpretar como la suma de volúmenes por encima y por debajo del plano $z = 0$. Los volúmenes por encima de $z = 0$ se cuentan como positivos y los por debajo como negativos. El teorema de Fubini 12.3 según se enunció sigue siendo válido en el caso de que $f(x, y)$ sea negativo o cambie de signo en R ; esto es, no hay restricciones en el signo de f en las hipótesis del teorema.

Ejemplo. Sea R el rectángulo $[-2, 1] \times [0, 1]$ y sea f definida por $f(x, y) = y(x^3 - 12x)$. Calcula la integral

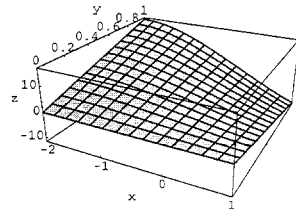
$$\int_R f(x, y) \, dx \, dy = \int_R y(x^3 - 12x) \, dx \, dy.$$

Solución:

Esta función toma valores positivos y negativos en el rectángulo R , como se puede ver en la gráfica.

Por el teorema de Fubini,

$$\begin{aligned} \int_R f(x, y) \, dx \, dy &= \int_0^1 \int_{-2}^1 y(x^3 - 12x) \, dx \, dy = \\ \int_0^1 \left[\int_{-2}^1 y(x^3 - 12x) \, dx \right] dy &= \frac{57}{4} \int_0^1 y \, dy = \frac{57}{8}. \end{aligned}$$



De manera alternativa, integrando primero respecto a y , hallamos

$$\begin{aligned} \int_R y(x^3 - 12x) \, dx \, dy &= \int_{-2}^1 \left[\int_0^1 y(x^3 - 12x) \, dy \right] dx = \\ &= \frac{1}{2} \int_{-2}^1 (x^3 - 12x) \, dx = \frac{1}{2} \left[\frac{x^4}{4} - 6x^2 \right]_{-2}^1 = \frac{57}{8}. \end{aligned}$$

Ejercicios.

1) Halla las sumas inferior y superior de la función $f(x, y) = x + 2y$ relativas a la partición $P = P_1 \times P_2$ de $R = \{(x, y) / 0 \leq x \leq 2, 0 \leq y \leq 1\}$, si

a) $P_1 = \{0, 1, \frac{3}{2}, 2\}$ y $P_2 = \{0, \frac{1}{2}, 1\}$;

b) $P_1 = \{0, \frac{1}{2}, 1, \frac{3}{2}, 2\}$ y $P_2 = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$.

2) Halla las sumas inferior y superior de la función $f(x, y) = x - y$ relativas a la partición $P = P_1 \times P_2$ de $R = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq 1\}$, si

a) $P_1 = \{0, \frac{1}{2}, \frac{3}{4}, 1\}$ y $P_2 = \{0, \frac{1}{2}, 1\}$;

b) $P_1 = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ y $P_2 = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$.

3) Evalúa las integrales siguientes, donde R es el rectángulo $[0, 1] \times [0, 1]$.

a) $\int_R (x^3 + y^2) dA$;

b) $\int_R (y e^{xy}) dA$;

c) $\int_R (x^2 y^2 \cos x^3) dA$;

d) $\int_R \log[(x+1)(y+1)] dA$;

e) $\int_R (ax + by + c) dx dy$;

f) $\int_R (x^m y^n) dx dy, m, n > 0$;

g) $\int_R (x^2 + 2xy + y\sqrt{x}) dx dy$;

h) $\int_R \sin(x+y) dx dy$.

4) Halla el volumen acotado por la gráfica de la función $f(x, y)$ en el rectángulo R indicado y los cuatro lados verticales del rectángulo R .

a) $f(x, y) = x^2 + y^4, R = [0, 1] \times [0, 1]$;

b) $f(x, y) = \text{sen } y, R = [0, 1] \times [0, \pi/2]$;

c) $f(x, y) = x^2 + y, R = [0, 1] \times [1, 2]$.

5) Calcula las siguientes integrales dobles por integración sucesiva.

a) $\iint_R (x^3 + 3x^2y + y^3) dA$, donde $R = [0, 1] \times [0, 1]$;

b) $\iint_R (\sqrt{y} + x - 3xy^2) dA$, donde $R = [0, 1] \times [1, 3]$;

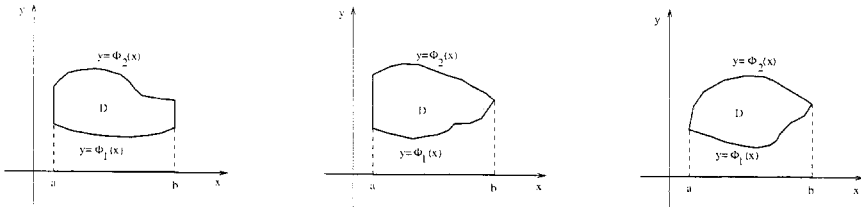
c) $\iint_R (\text{sen}(x+y)) dx dy$, donde $R = [0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]$;

d) $\iint_R |\cos(x+y)| dx dy$, donde $R = [0, \pi] \times [0, \pi]$.

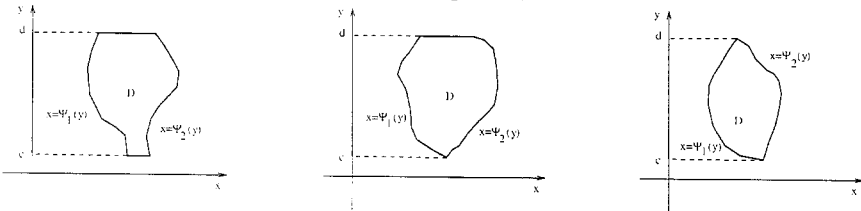
12.3 La integral doble sobre regiones más generales.

En esta sección vamos a definir la integral doble $\int_D f(x, y) dA$ en regiones D más generales que rectángulos y desarrollaremos una técnica para evaluar este tipo de integrales. A este fin, definimos tres tipos especiales de subconjuntos D del plano xy , que llamaremos **regiones elementales**, y cuya frontera ∂D es del tipo de conjunto de discontinuidades de una función permitida en el teorema 12.2.

Definición. Sean $\Phi_1: [a, b] \rightarrow \mathbb{R}$ y $\Phi_2: [a, b] \rightarrow \mathbb{R}$ dos funciones continuas con valores reales que satisfagan $\Phi_1(x) \leq \Phi_2(x)$ para todo $x \in [a, b]$. Se dirá **región de tipo 1** al conjunto D dado por $D = \{(x, y) / x \in [a, b], \Phi_1(x) \leq y \leq \Phi_2(x)\}$. Las curvas y segmentos de rectas que acotan la región D , forman juntos la frontera ∂D de D .

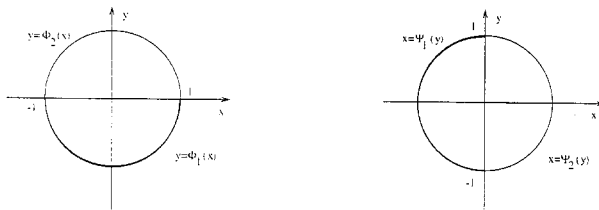


Definición. Decimos que una **región D es de tipo 2** si existen funciones continuas $\Psi_1: [c, d] \rightarrow \mathbb{R}$ y $\Psi_2: [c, d] \rightarrow \mathbb{R}$ que satisfagan $\Psi_1(y) \leq \Psi_2(y)$ para todo $y \in [c, d]$, y el conjunto D viene dado por $D = \{(x, y) / y \in [c, d], \Psi_1(y) \leq x \leq \Psi_2(y)\}$. De nuevo las curvas y segmentos de rectas que acotan la región D , forman su frontera ∂D .



Finalmente una **región de tipo 3** es aquella que es del tipo 1 y del tipo 2 simultáneamente; es decir, se puede describir tanto como una región de tipo 1 como una región de tipo 2. Un ejemplo de región de tipo 3 es el disco unitario para el cual

$$\Phi_1(x) = -\sqrt{1-x^2}, \quad \Phi_2(x) = \sqrt{1-x^2}, \quad \Psi_1(y) = -\sqrt{1-y^2}, \quad \Psi_2(y) = \sqrt{1-y^2}.$$



Para poder definir integrales dobles en regiones elementales del plano xy , realizamos lo siguiente. Sea D es una región elemental en el plano, entonces escogemos un rectángulo R que contenga a D . Dada $f: D \rightarrow \mathbb{R}$ continua (y por lo tanto acotada), extendemos la función f a una función f^* definida en todo el rectángulo R como

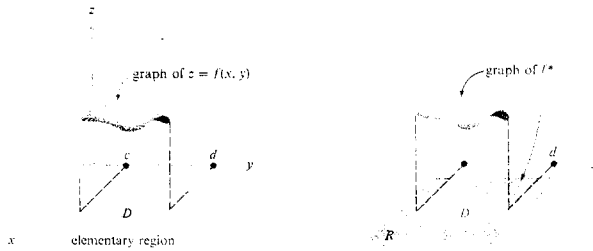
$$f^*(x, y) = \begin{cases} f(x, y) & (x, y) \in D \\ 0 & (x, y) \notin D. \quad (x, y) \in R \end{cases}$$

Dado que la función f está acotada, también la nueva función f^* lo está: además f^* es continua, excepto quizá en la frontera ∂D de D . Sin embargo, la frontera ∂D está formada por gráficas de funciones continuas, de modo que f^* es integrable sobre R por el teorema 12.2. Por lo tanto

Definición. Definimos $\int_D f(x, y) dA$, la **integral de f sobre el conjunto D** como la integral sobre el rectángulo R de la función f^* :

$$\int_D f(x, y) dA = \int_R f^*(x, y) dA.$$

Cuando $f(x, y) \geq 0$ en D , podemos interpretar la integral doble como el volumen de la región tridimensional entre la gráfica de f y D .



Teorema 12.5 Sea $D = \{(x, y) / x \in [a, b], \Phi_1(x) \leq y \leq \Phi_2(x)\}$ una región del tipo 1. entonces

$$\int_D f(x, y) dA = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} f(x, y) dy dx. \tag{12.4}$$

Demostración:

Si $R = [a, b] \times [c, d]$ es un rectángulo cualquiera que contiene al dominio D , podemos usar los resultados sobre integrales iteradas de la sección 12.2 para obtener

$$\int_D f(x, y) dA = \int_R f^*(x, y) dA = \int_a^b \int_c^d f^*(x, y) dy dx = \int_c^d \int_a^b f^*(x, y) dx dy$$

donde f^* es igual a f en D y cero fuera de D . Si D es una región de tipo 1, las funciones $\Phi_1(x)$ y $\Phi_2(x)$ satisfacen $\Phi_1(x) \leq \Phi_2(x)$ para toda $x \in [a, b]$. Podemos considerar entonces

12.3. LA INTEGRAL DOBLE SOBRE REGIONES MÁS GENERALES.

la integral iterada $\int_a^b \int_c^d f^*(x, y) dy dx$, y en particular la integral interior $\int_c^d f^*(x, y) dy$ para deducir que como por definición es $f^*(x, y) = 0$ si $y < \Phi_1(x)$ o $y > \Phi_2(x)$, luego

$$\int_c^d f^*(x, y) dy = \int_{\Phi_1(x)}^{\Phi_2(x)} f^*(x, y) dy = \int_{\Phi_1(x)}^{\Phi_2(x)} f(x, y) dy.$$

Finalmente, sustituyendo en la expresión anterior

$$\int_D f(x, y) dA = \int_a^b \int_c^d f^*(x, y) dy dx = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} f(x, y) dy dx.$$

c.q.d

En el caso $f(x, y) = 1$ para todo $(x, y) \in D$, la integral $\int_D f(x, y) dA$ es el área de D . De la fórmula (12.4) obtenemos

$$\int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} f(x, y) dy dx = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} dy dx = \int_a^b [\Phi_2(x) - \Phi_1(x)] dx = A(D).$$

Nótese que este resultado para el área es consistente con los resultados de cálculo de una variable para el área de una región entre dos curvas.

Con una demostración parecida a la del teorema 12.5 se puede demostrar lo siguiente.

Teorema 12.6 *Sea $D = \{(x, y) / y \in [c, d], \Psi_1(y) \leq x \leq \Psi_2(y)\}$ una región del tipo 2, entonces*

$$\int_D f(x, y) dA = \int_c^d \int_{\Psi_1(y)}^{\Psi_2(y)} f(x, y) dx dy. \quad (12.5)$$

En el caso $f(x, y) = 1$ para todo $(x, y) \in D$, la integral $\int_D f(x, y) dA$ es el área de D . De la fórmula (12.5) obtenemos

$$\int_c^d \int_{\Psi_1(y)}^{\Psi_2(y)} f(x, y) dx dy = \int_c^d \int_{\Psi_1(y)}^{\Psi_2(y)} dx dy = \int_c^d [\Psi_2(y) - \Psi_1(y)] dy = A(D).$$

Es interesante notar que estas integrales no dependen de la selección del rectángulo R . Además, para integrales dobles sobre regiones más generales valen las mismas propiedades vistas en la sección 12.2.

Como propiedad adicional enunciamos, sin demostración, el siguiente teorema:

Teorema 12.7 Teorema del valor medio para integrales dobles.

Sean f y g dos funciones continuas en una región $D \subset \mathbb{R}^2$. Si g es no negativa en D , existe un punto (x_0, y_0) en D para el cual

$$\int_D f(x, y) g(x, y) dx dy = f(x_0, y_0) \int_D g(x, y) dx dy.$$

Llamaremos al valor $f(x_0, y_0)$ la **media ponderada por g de f en D** .

Si escogemos $g(x, y) = 1$ para todo $(x, y) \in D$ la fórmula anterior se reduce a

$$\int_D f(x, y) dx dy = f(x_0, y_0) \int_D dx dy = f(x_0, y_0) \cdot (\text{Area } (D)).$$

Diremos que $f(x_0, y_0)$ es el **valor medio de f en D** .

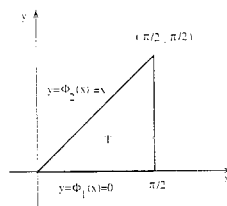
Ejemplo 1. Sea T el triángulo $T = \{(x, y) / 0 \leq x \leq \pi/2, 0 \leq y \leq x\}$; calcula la siguiente integral doble

$$\int_T (x^3 y + \cos x) dA.$$

Solución:

Considerando la región T como de tipo 1 y usando la fórmula (12.4), tenemos

$$\begin{aligned} \int_T (x^3 y + \cos x) dA &= \int_0^{\pi/2} \int_0^x (x^3 y + \cos x) dy dx = \\ \int_0^{\pi/2} \left[x^3 \frac{y^2}{2} + y \cos x \right]_0^x dx &= \int_0^{\pi/2} \left(\frac{x^5}{2} + x \cos x \right) dx = \\ \left[\frac{x^6}{12} + x \operatorname{sen} x + \cos x \right]_0^{\pi/2} &= \frac{\pi^6}{768} + \frac{\pi}{2} - 1. \end{aligned}$$



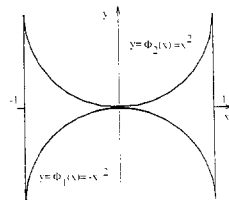
Ejemplo 2. Sea D el dominio $D = \{(x, y) / -1 \leq x \leq 1, -x^2 \leq y \leq x^2\}$; calcula la integral doble

$$\int_D (x^2 - y) dA.$$

Solución:

Por su propia definición, la región D es una región de tipo 1. Usando la fórmula (12.4), tenemos

$$\begin{aligned} \int_D (x^2 - y) dA &= \int_{-1}^1 \int_{-x^2}^{x^2} (x^2 - y) dy dx = \\ \int_{-1}^1 \left[x^2 y - \frac{y^2}{2} \right]_{-x^2}^{x^2} dx &= \int_{-1}^1 2x^4 dx = \\ \left[\frac{2}{5} x^5 \right]_{-1}^1 &= \frac{4}{5}. \end{aligned}$$



Ejemplo 3. Halla el volumen del tetraedro acotado por los planos $y = 0$, $z = 0$, $x = 0$ e $y - x + z = 1$.

Solución:

El tetraedro dado tiene una base triangular D cuyos puntos (x, y) satisfacen

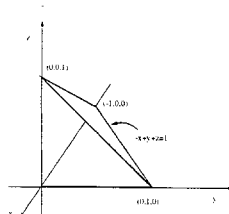
$$-1 \leq x \leq 0 \quad \text{y} \quad 0 \leq y \leq 1 + x;$$

12.3. LA INTEGRAL DOBLE SOBRE REGIONES MÁS GENERALES.

por lo tanto podemos ver D como una región de tipo 1.

Para cualquier punto (x, y) en D , la altura de la superficie z sobre (x, y) es $1+x-y$. Así, el volumen que buscamos está dado por la integral

$$\int_D (1+x-y) dA.$$



Usando la fórmula (12.4), obtenemos

$$\begin{aligned} \int_D (1+x-y) dA &= \int_{-1}^0 \int_0^{1+x} (1+x-y) dy dx = \int_{-1}^0 \left[(1+x)y - \frac{y^2}{2} \right]_0^{1+x} dx = \\ &= \int_{-1}^0 \left[\frac{(1+x)^2}{2} \right] dx = \left[\frac{(1+x)^3}{6} \right]_{-1}^0 = \frac{1}{6}. \end{aligned}$$

Ejercicios.

1) Dibuja las regiones de integración y calcula las siguientes integrales dobles:

- $\iint_S x \cos(x+y) dx dy$, S triángulo de vértices $(0,0)$, $(\pi,0)$, (π,π) ;
- $\iint_S (1+x) \sin y dx dy$, S trapecoide de vértices $(0,0)$, $(1,0)$, $(1,2)$, $(0,1)$;
- $\iint_D (x^2 - y^2) dx dy$, D región limitada por $y = \sin x$ y el intervalo $[0, \pi]$.

2) Calcula por doble integración el volumen del conjunto de ordenadas de f sobre S :

- $f(x,y) = x^2 + y^2$, y $S = \{(x,y) / |x| \leq 1, |y| \leq 1\}$;
- $f(x,y) = 3x + y$, y $S = \{(x,y) / 4x^2 + 9y^2 \leq 36, x > 0, y > 0\}$;
- $f(x,y) = y + 2x + 20$, y $S = \{(x,y) / x^2 + y^2 \leq 16\}$.

3) Evalúa las siguientes integrales iteradas y traza las regiones D .

- $\int_0^1 \int_0^{x^2} dy dx$;
- $\int_1^2 \int_{2x}^{3x+1} dy dx$;
- $\int_0^1 \int_1^{e^x} (x+y) dy dx$;
- $\int_0^1 \int_{x^3}^{x^2} y dy dx$;
- $\int_{-3}^2 \int_0^{y^2} (x^2 + y) dx dy$;
- $\int_0^{\pi/2} \int_0^{\cos x} (y \sin x) dy dx$.

4) Halla el área de la región del primer cuadrante delimitada por $xy = 2$, $y = 1$, $y = x + 1$.

5) Calcula por integración doble el área de la región acotada por los siguientes pares de curvas:

a) $x^2 = 4y, \quad 2y - x - 4 = 0;$

b) $y = x. \quad x = 4y - y^2;$

c) $y = x, \quad 4y^3 = x^2;$

d) $x + y = 5. \quad xy = 6.$

6) Sea D la región acotada por el eje y y la parábola $x = -4y^2 + 3$. Calcula la integral doble $\int_D x^3 y \, dx \, dy$.

7) Halla el volumen del sólido limitado superiormente por $z = x + y$ e inferiormente por el cuadrado unidad $0 \leq x \leq 1, 0 \leq y \leq 1$.

12.4 Cambio del orden de integración.

Si el conjunto D es una región de tipo 3, quiere decir que es del tipo 1 y del tipo 2 a la vez, entonces podrá expresarse como el conjunto de puntos (x, y) tales que

$$a \leq x \leq b, \quad \Phi_1(x) \leq y \leq \Phi_2(x)$$

y también como el conjunto de puntos (x, y) tales que

$$c \leq y \leq d, \quad \Psi_1(y) \leq x \leq \Psi_2(y).$$

Por lo tanto tenemos las fórmulas

$$\int_D f(x, y) \, dA = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} f(x, y) \, dy \, dx = \int_c^d \int_{\Psi_1(y)}^{\Psi_2(y)} f(x, y) \, dx \, dy. \quad (12.6)$$

Si tenemos que calcular una de las integrales iteradas, lo podemos hacer evaluando la otra; esta técnica se llama **cambio del orden de integración**. Suele ser útil realizar estos cambios al evaluar integrales iteradas, pues quizá una de las mismas sea más difícil de calcular que la otra.

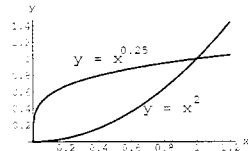
Ejemplo 1. Calcula

$$\int_D (\sqrt{x} - y^2) \, dA$$

con D el dominio de intersección de las curvas $y = x^2$ e $y = \sqrt[3]{x}$.

Solución:

Al proyectar D en el eje x obtenemos el intervalo cerrado $[0, 1]$. Luego el dominio D consiste en $\{(x, y) / 0 \leq x \leq 1, x^2 \leq y \leq \sqrt[3]{x}\}$, que es una región de tipo 1. Usando la fórmula (12.4), que coincide con la primera de (12.6), obtenemos



12.4. CAMBIO DEL ORDEN DE INTEGRACIÓN.

$$\begin{aligned} \int_D (\sqrt{x} - y^2) dA &= \int_0^1 \int_{x^2}^{\sqrt[3]{x}} (\sqrt{x} - y^2) dy dx = \\ &= \int_0^1 \left[y\sqrt{x} - \frac{y^3}{3} \right]_{x^2}^{\sqrt[3]{x}} dx = \int_0^1 \left(\frac{2}{3} \sqrt[3]{x^3} - \sqrt{x^5} + \frac{1}{3} x^6 \right) dx = \\ &= \left[\frac{8}{21} \sqrt[3]{x^7} - \frac{2}{7} \sqrt{x^7} + \frac{1}{21} x^7 \right]_0^1 = \frac{1}{7}. \end{aligned}$$

También podemos integrar en el orden inverso. Al proyectar D en el eje y obtenemos el intervalo cerrado $[0, 1]$. Luego el dominio $D = \{(x, y) / y^4 \leq x \leq \sqrt{y}, 0 \leq y \leq 1\}$ es una región de tipo 2. Usando la fórmula (12.5), que coincide con la segunda de (12.6), obtenemos

$$\begin{aligned} \int_D (\sqrt{x} - y^2) dA &= \int_0^1 \int_{y^4}^{\sqrt{y}} (\sqrt{x} - y^2) dx dy = \int_0^1 \left[\frac{2}{3} \sqrt{x^3} - y^2 x \right]_{y^4}^{\sqrt{y}} dy = \\ &= \int_0^1 \left(\frac{2}{3} \sqrt[3]{y^3} - \sqrt{y^5} + \frac{1}{3} y^6 \right) dy = \left[\frac{8}{21} \sqrt[3]{y^7} - \frac{2}{7} \sqrt{y^7} + \frac{1}{21} y^7 \right]_0^1 = \frac{1}{7}. \end{aligned}$$

Ejemplo 2. Calcula mediante integración doble el área de la región D comprendida entre las curvas $\sqrt{x} + \sqrt{y} = a$ y $x + y = a$.

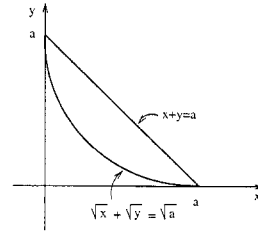
Solución:

La región D está representada en la figura. Su área viene dada por la integral doble

$$\int_D dx dy.$$

Podemos proyectar D sobre el eje x y escribir las fronteras como funciones de x :

$$y = (\sqrt{a} - \sqrt{x})^2 \quad \text{e} \quad y = a - x.$$



Luego el dominio D consiste en $\{(x, y) / 0 \leq x \leq a, (\sqrt{a} - \sqrt{x})^2 \leq y \leq a - x\}$, que es una región de tipo 1. Usando la primera parte de la fórmula (12.6), obtenemos

$$\begin{aligned} \int_D dx dy &= \int_0^a \int_{(\sqrt{a}-\sqrt{x})^2}^{a-x} dy dx = \int_0^a \left[(a-x) - (\sqrt{a} - \sqrt{x})^2 \right] dx = \\ &= \int_0^a (-2x + 2\sqrt{a}\sqrt{x}) dx = \left[-x^2 + \frac{4}{3} \sqrt{a}\sqrt{x^3} \right]_0^a = \frac{1}{3} a^2. \end{aligned}$$

También podemos proyectar D sobre el eje y y escribir las fronteras como funciones de y

$$x = (\sqrt{a} - \sqrt{y})^2 \quad \text{y} \quad x = a - y.$$

Luego podemos escribir $D = \{(x, y) / (\sqrt{a} - \sqrt{y})^2 \leq x \leq a - y, 0 \leq y \leq a\}$, que es una región de tipo 2. Usando la segunda parte de la fórmula (12.6), obtenemos

$$\begin{aligned} \int_D dy dx &= \int_0^a \int_{(\sqrt{a}-\sqrt{y})^2}^{a-y} dx dy = \int_0^a [(a-y) - (\sqrt{a}-\sqrt{y})^2] dy = \\ &= \int_0^a (-2y + 2\sqrt{a}\sqrt{y}) dy = \left[-y^2 + \frac{4}{3}\sqrt{a}\sqrt{y^3}\right]_0^a = \frac{1}{3}a^2. \end{aligned}$$

Ejemplo 3. Evalúa la siguiente integral cambiando el orden de integración:

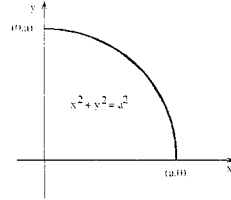
$$\int_0^a \int_0^{\sqrt{a^2-x^2}} \sqrt{a^2-x^2} dy dx.$$

Solución:

Por como viene dada la integral, el dominio de integración D está formado por los puntos (x, y) tales que

$$0 \leq x \leq a \quad \text{y} \quad 0 \leq y \leq \sqrt{a^2-x^2},$$

es decir, es el primer cuadrante del disco de radio a y centro en el origen.



Por lo tanto D también puede describirse como el conjunto de puntos (x, y) que satisface $0 \leq y \leq a$ y $0 \leq x \leq \sqrt{a^2-y^2}$. Usando la fórmula (12.6), obtenemos

$$\begin{aligned} \int_0^a \int_0^{\sqrt{a^2-x^2}} \sqrt{a^2-x^2} dy dx &= \int_0^a \int_0^{\sqrt{a^2-y^2}} \sqrt{a^2-y^2} dx dy = \\ \int_0^a [x\sqrt{a^2-y^2}]_0^{\sqrt{a^2-y^2}} dy &= \int_0^a (a^2-y^2) dy = \left[a^2y - \frac{y^3}{3}\right]_0^a = \frac{2}{3}a^3. \end{aligned}$$

Calcular la primera integral iterada directamente es posible pero notablemente más complicado.

El siguiente ejemplo muestra que a veces puede ser *imposible* evaluar una integral iterada y, sin embargo, ser posible evaluar la integral obtenida cambiando el orden de integración.

Ejemplo. Evalúa

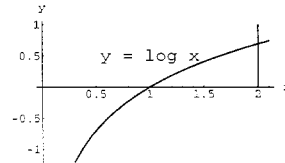
$$\int_1^2 \int_0^{\log x} (x-1)\sqrt{1+e^{2y}} dy dx.$$

Solución:

Por como viene dada la integral, el dominio de integración D está formado por los puntos (x, y) tales que

$$1 \leq x \leq 2 \quad \text{y} \quad 0 \leq y \leq \log x,$$

es decir, viene dado como recinto de tipo 1.



12.4. CAMBIO DEL ORDEN DE INTEGRACIÓN.

Sin embargo, D es de tipo 3 y por lo tanto puede expresarse también como

$$0 \leq y \leq \log(2) \quad \text{y} \quad e^y \leq x \leq 2.$$

Así, usando la fórmula (12.6), la integral iterada pedida es igual a

$$\begin{aligned} \int_1^2 \int_0^{\log x} (x-1) \sqrt{1+e^{2y}} \, dy \, dx &= \int_0^{\log(2)} \sqrt{1+e^{2y}} \left[\int_{e^y}^2 (x-1) \, dx \right] \, dy = \\ &= \int_0^{\log(2)} \sqrt{1+e^{2y}} \left[\frac{x^2}{2} - x \right]_{e^y}^2 \, dy = \int_0^{\log(2)} \sqrt{1+e^{2y}} \left(e^y - \frac{e^{2y}}{2} \right) \, dy. \end{aligned}$$

Esta última integral es suma de dos términos. Si en el primero realizamos el cambio de variable $v = e^y$ y en el segundo $u = e^{2y}$ llegamos a

$$\begin{aligned} \int_1^2 \sqrt{1+v^2} \, dv - \frac{1}{4} \int_1^4 \sqrt{1+u} \, du &= \frac{1}{2} \left[v \sqrt{1+v} + \log(\sqrt{1+v} + v) \right]_1^2 - \\ &- \left[\frac{1}{6} \sqrt{(1+u)^3} \right]_1^4 = \frac{1}{2} \left(2\sqrt{5} - \sqrt{2} + \log \frac{\sqrt{5}+2}{\sqrt{2}+1} \right) - \frac{1}{6} (5\sqrt{5} - 2\sqrt{2}). \end{aligned}$$

Concluimos esta sección con algunas consideraciones sobre la **simetría en la integración doble**.

En el caso de funciones de una variable, teníamos que si f es continua sobre un intervalo que es simétrico respecto del origen, digamos $[-a, a]$,

$$\text{si } f \text{ es impar} \quad \Rightarrow \quad \int_{-a}^a f(x) \, dx = 0;$$

$$\text{si } f \text{ es par} \quad \Rightarrow \quad \int_{-a}^a f(x) \, dx = 2 \int_0^a f(x) \, dx.$$

Para las integrales dobles tenemos resultados similares.

Sean D un recinto simétrico respecto del eje y y D_d la mitad derecha, luego

$$\text{si } f \text{ es impar en } x [f(-x, y) = -f(x, y)] \quad \Rightarrow \quad \int_D f(x, y) \, dx \, dy = 0;$$

$$\text{si } f \text{ es par en } x [f(-x, y) = f(x, y)] \quad \Rightarrow \quad \int_D f(x, y) \, dx \, dy = 2 \int_{D_d} f(x, y) \, dx \, dy.$$

Sea ahora D un recinto simétrico respecto del eje x , y D_s la mitad superior, luego

$$\text{si } f \text{ es impar en } y [f(x, -y) = -f(x, y)] \quad \Rightarrow \quad \int_D f(x, y) \, dx \, dy = 0;$$

$$\text{si } f \text{ es par en } y [f(x, -y) = f(x, y)] \quad \Rightarrow \quad \int_D f(x, y) \, dx \, dy = 2 \int_{D_s} f(x, y) \, dx \, dy.$$

Ejemplo 1. Sea D la región finita comprendida entre las dos rectas $y = x$, $y = -x$, para $-a \leq x \leq a$. Calcula

$$\int_D (2x - \operatorname{sen} x^2 y) \, dx \, dy.$$

Solución:

La simetría respecto del eje y nos da

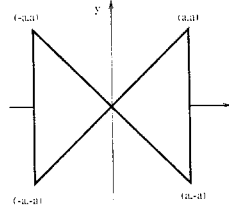
$$\int_D 2x \, dx \, dy = 0,$$

siendo el integrando impar en x . De la misma manera, la simetría respecto del eje x nos da

$$\int_D \operatorname{sen} x^2 y \, dx \, dy = 0.$$

siendo el integrando impar en y . Luego (comprobado evaluando la integral)

$$\int_D (2x - \operatorname{sen} x^2 y) \, dx \, dy = 0.$$



Ejemplo 2. Calculemos otra vez la integral doble

$$\int_D (x^2 - y) \, dA,$$

donde D es el dominio $D = \{(x, y) / -1 \leq x \leq 1, -x^2 \leq y \leq x^2\}$.

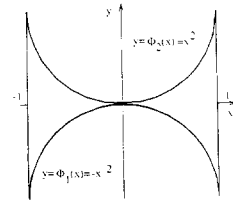
Solución:

Ya vimos que la solución es $\frac{4}{5}$.

Ahora utilizemos la simetría del recinto D .

Por simetría de la función $f(x, y) = y$ respecto del eje x , tenemos

$$\int_D y \, dx \, dy = 0.$$



Por otro lado la función $f(x, y) = x^2$ resulta ser simétrica respecto del eje x y también respecto del eje y ; luego

$$\int_D x^2 \, dx \, dy = 2 \int_{D_d} x^2 \, dx \, dy = 4 \int_{(D_d)_s} x^2 \, dx \, dy$$

donde $(D_d)_s$ es la mitad superior de la mitad derecha de D . Finalmente

$$\int_D (x^2 - y) \, dA = 4 \int_{(D_d)_s} x^2 \, dx \, dy = 4 \int_0^1 \int_0^{x^2} x^2 \, dy \, dx = 4 \int_0^1 x^4 \, dx = \frac{4}{5}.$$

12.4. CAMBIO DEL ORDEN DE INTEGRACIÓN.

Ejemplo 3. Calcula el volumen de la porción del cilindro $x^2 + y^2 = b^2$ comprendida entre los planos $y + z = a^2$ y $z = 0$, con $a^2 \geq b > 0$.

Solución:

El sólido en cuestión está acotado inferiormente por el disco

$$D = \{(x, y) / 0 \leq x^2 + y^2 \leq b^2\},$$

y superiormente, por el plano

$$z = a^2 - y.$$

Entonces el volumen viene dado por la integral

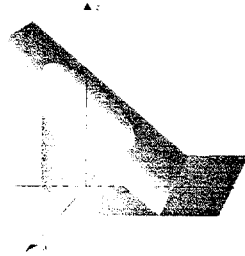
$$\int_D (a^2 - y) dx dy = 0.$$

Dado que D es simétrico respecto del eje x tenemos

$$\int_D y dx dy = 0.$$

Luego (comprobadlo evaluando la integral)

$$\int_D (a^2 - y) dx dy = \int_D a^2 dx dy = a^2 \int_D dx dy = a^2 (\text{área de } D) = \pi a^2 b^2.$$



Ejercicios.

- 1) Escribe las ecuaciones correspondientes a los límites de los recintos en que están definidas las siguientes integrales dobles. Dibuja estos recintos. Inverte el orden de integración en las siguientes integrales dobles.

a) $\int_0^4 \int_{3x^2}^{12x} f(x, y) dy dx;$

b) $\int_0^1 \int_{y^2/2}^{3-y^2} f(x, y) dx dy;$

c) $\int_{a/2}^a \int_0^{(2ax-x^2)^{1/2}} f(x, y) dy dx;$

d) $\int_0^\pi \int_0^{\sin x} f(x, y) dy dx;$

e) $\int_1^3 \int_{x/3}^{2x} f(x, y) dy dx;$

f) $\int_1^3 \int_{x^2}^{x+9} f(x, y) dy dx;$

g) $\int_0^4 \int_y^{10-y} f(x, y) dx dy;$

h) $\int_{-1}^2 \int_{x^2}^{x+2} f(x, y) dy dx;$

i) $\int_0^1 \int_{x^4}^{x^2} f(x, y) dy dx;$

j) $\int_0^1 \int_0^{y^2} f(x, y) dx dy;$

k) $\int_0^1 \int_{-y}^y f(x, y) dx dy;$

l) $\int_{1/2}^1 \int_{x^3}^x f(x, y) dy dx;$

m) $\int_1^4 \int_x^{2x} f(x, y) dy dx;$

n) $\int_{-x}^3 \int_{-x}^{x^2} f(x, y) dy dx.$

2) Escribe las ecuaciones correspondientes a los límites de los recintos en los que están definidas las siguientes integrales dobles. Dibuja estos recintos. Inverte el orden de integración en las siguientes integrales dobles y calcula las integrales de la dos maneras.

a) $\int_0^1 \int_x^1 x y \, dy \, dx;$
 b) $\int_0^1 \int_1^{2-y} (x+y)^2 \, dx \, dy;$
 c) $\int_{-3}^1 \int_{-\sqrt{9-y^2}}^{\sqrt{9-y^2}} x^2 \, dx \, dy.$

3) Evalúa las siguientes integrales dobles.

a) $\int_0^1 \int_0^{3y} e^{x+y} \, dx \, dy;$ b) $\int_1^4 \int_1^{\sqrt{x}} (x^2 + y^2) \, dy \, dx;$
 c) $\int_0^1 \int_{1-y}^1 (x+y^2) \, dx \, dy;$ d) $\int_{-1}^0 \int_{-\sqrt{y+1}}^{\sqrt{y+1}} x^2 \, dx \, dy;$
 e) $\iint_D y^2 \sqrt{x} \, dx \, dy, \quad D = \{(x, y) / x > 0, y > x^2, y < 10 - x^2\};$
 f) $\iint_D e^{x-y} \, dx \, dy, \quad D = \{\text{triángulo con vértices } (0, 0), (1, 3) \text{ y } (2, 2)\};$
 g) $\iint_D y^3 (x^2 + y^2)^{-3/2} \, dx \, dy, \quad D = \{(x, y) / \frac{1}{2} \leq y \leq 1, x^2 + y^2 \leq 1\};$
 h) $\iint_D (x^2 + 2xy + 2) \, dx \, dy, \quad D$ es la región acotada por el eje x , las rectas $x = 0$ y $x = 2$ y la gráfica de $y = -x^2 + x$;
 i) $\iint_D (x^2 + 3y^3) \, dx \, dy, \quad D = \{(x, y) / 0 \leq x^2 + y^2 \leq 1\};$
 j) $\iint_D \sqrt{xy} \, dx \, dy, \quad D = \{(x, y) / 0 \leq y \leq 1, y^2 \leq x \leq y\};$
 k) $\iint_D y e^x \, dx \, dy, \quad D = \{(x, y) / 0 \leq y \leq 1, 0 \leq x \leq y^2\};$
 l) $\iint_D (x+y) \, dx \, dy$, donde D es la región acotada entre $y = x^3$ e $y = x^4$, con $x \in [-1, 1]$.

4) Calcula por integración doble el área de la región acotada por los siguientes pares de curvas.

a) $x^2 = 4y, \quad 2y - x - 4 = 0;$ b) $y = x, \quad x = 4y - y^2;$
 c) $y = x, \quad 4y^3 = x^2;$ d) $x + y = 5, \quad xy = 6.$

12.5 La integral triple.

Una vez familiarizados con las integrales dobles $\iint_D f(x, y) dx dy$, resultará fácil entender las integrales triples

$$\iiint_T f(x, y, z) dx dy dz.$$

Básicamente, la única diferencia será que en lugar de trabajar con funciones de dos variables continuas en una región D del plano, trabajaremos con funciones de tres variables continuas en una determinada porción T del espacio tridimensional.

Dada una función continua $f: R \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, donde R es algún paralelepípedo rectangular en \mathbb{R}^3 , podemos definir la integral de f sobre R como límite de sumas, así como lo hicimos para una función de dos variables.

El paralelepípedo rectangular $R \subset \mathbb{R}^3$ es el producto cartesiano $R = [a, b] \times [c, d] \times [e, f]$. **Definición.** Por una **partición regular** P de R de orden n , entendemos tres colecciones ordenadas de $n + 1$ puntos igualmente espaciados $\{x_i\}_{i=0}^n$, $\{y_j\}_{j=0}^n$ y $\{z_k\}_{k=0}^n$; esto es, puntos que satisfacen

$$a = x_0 < x_1 < x_2 < \dots < x_n = b,$$

$$c = y_0 < y_1 < y_2 < \dots < y_n = d$$

$$e = z_0 < z_1 < z_2 < \dots < z_n = f$$

y

$$\Delta x = x_{i+1} - x_i = \frac{b-a}{n},$$

$$\Delta y = y_{j+1} - y_j = \frac{d-c}{n},$$

$$\Delta z = z_{k+1} - z_k = \frac{f-e}{n}.$$

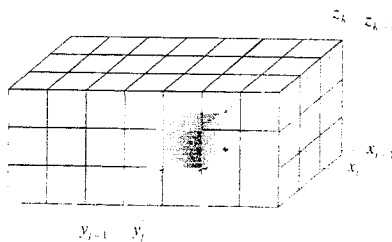
Sea R_{ijk} el ijk -ésimo paralelepípedo rectangular $[x_i, x_{i+1}] \times [y_j, y_{j+1}] \times [z_k, z_{k+1}]$ y \bar{c}_{ijk} cualquier punto en R_{ijk} . Supongamos que $f: R \rightarrow \mathbb{R}$ es una función acotada con valores reales. Podemos entonces formar la suma

$$S_n = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} f(\bar{c}_{ijk}) \Delta x \Delta y \Delta z = \sum_{i,j,k=0}^{n-1} f(\bar{c}_{ijk}) \Delta V,$$

donde $\Delta V = \Delta x \Delta y \Delta z$. Esta suma de Riemann para f está tomada sobre todo índice i , j y k de 0 a $n - 1$, de modo que hay n^3 términos.

Definición. Sea f una función acotada de tres variables, definida en R . Si existe el límite $S = \lim_{n \rightarrow \infty} S_n$ (para cualquier selección de puntos \bar{c}_{ijk} en los paralelepípedos rectangulares R_{ijk}), lo llamaremos **integral triple de la función f sobre R** y lo denotamos por

$$\int_R f dV, \quad \int_R f(x, y, z) dV, \quad \int_R f(x, y, z) dx dy dz, \quad \iiint_R f(x, y, z) dx dy dz;$$



es decir

$$\iiint_R f(x, y, z) \, dx \, dy \, dz = \lim_{n \rightarrow \infty} S_n = S = \lim_{n \rightarrow \infty} \sum_{i,j,k=0}^{n-1} f(\bar{c}_{ijk}) \, \Delta x \, \Delta y \, \Delta z.$$

Como en el caso de la integral doble, podemos probar que las funciones continuas definidas en $R \subset \mathbb{R}^3$ son integrables. Más aún,

Teorema 12.8 *Las funciones acotadas cuyas discontinuidades están confinadas en gráficas de funciones continuas (tales como $x = \alpha(y, z)$, $y = \beta(x, z)$ o $z = \gamma(x, y)$) son integrables.*

Esto es el análogo del teorema 12.2.

También hay un teorema análogo al de Fubini 12.3 en el caso de integrales triples.

Teorema 12.9 Teorema de Fubini.

Sea $f: R \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ una función continua en el dominio $R = [a, b] \times [c, d] \times [e, f]$. Entonces las seis integrales iteradas posibles son iguales.

$$\begin{aligned} \int_a^b \int_c^d \int_e^f f(x, y, z) \, dz \, dy \, dx &= \int_e^f \int_c^d \int_a^b f(x, y, z) \, dx \, dy \, dz = \\ &= \int_a^b \int_e^f \int_c^d f(x, y, z) \, dy \, dz \, dx = \int_e^f \int_a^b \int_c^d f(x, y, z) \, dy \, dx \, dz = \\ &= \int_c^d \int_a^b \int_e^f f(x, y, z) \, dz \, dx \, dy = \int_c^d \int_e^f \int_a^b f(x, y, z) \, dx \, dz \, dy. \end{aligned} \quad (12.7)$$

En otras palabras, una integral triple se puede reducir a una triple integración iterada.

Para completar la analogía con la integral doble, consideremos el problema de evaluar integrales triples sobre conjuntos acotados más generales $W \subset \mathbb{R}^3$. Siendo el conjunto acotado, se puede encerrar en una caja R . Entonces extendemos la función f a una función f^* definida en todo el paralelepípedo rectangular R como

$$f^*(x, y, z) = \begin{cases} f(x, y, z) & (x, y, z) \in W \\ 0 & (x, y, z) \notin W. \end{cases} \quad (x, y, z) \in R$$

Dado que la función f está acotada, también la nueva función f^* lo está; además f^* es continua, excepto quizá en la frontera ∂W de W . Sin embargo, la frontera ∂W está formada por gráficas de funciones continuas, de modo que f^* es integrable sobre R por el teorema 12.8. Por lo tanto

Definición. Definimos $\int_D f(x, y, z) \, dV$, la **integral de f sobre el conjunto W** como la integral sobre el paralelepípedo rectangular R de la función f^* :

$$\int_W f(x, y, z) \, dV = \int_R f^*(x, y, z) \, dV.$$

12.5. LA INTEGRAL TRIPLE.

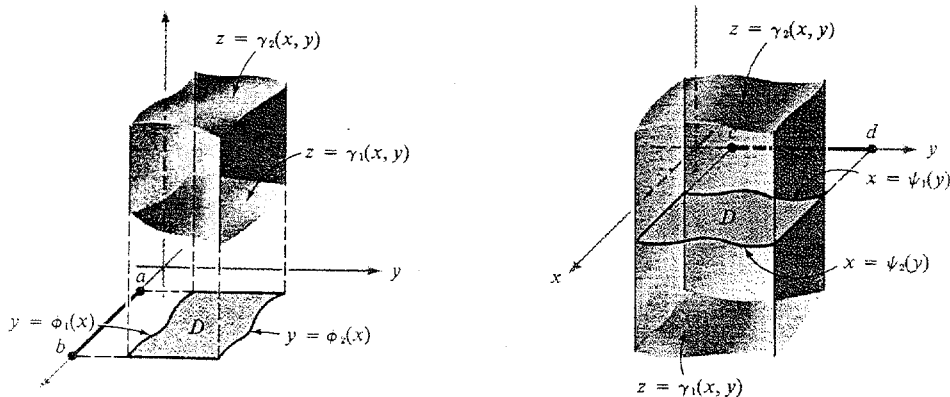
Como en el caso bidimensional, estas integrales son independientes de la selección del paralelepípedo rectangular R .

Nuevamente, como en el caso de dos variables, nos restringiremos a regiones de tipos especiales.

Definición. Una región W es del **tipo I** si se puede escribir como el conjunto de puntos (x, y, z) tales que

$$a \leq x \leq b, \quad \Phi_1(x) \leq y \leq \Phi_2(x), \quad \gamma_1(x, y) \leq z \leq \gamma_2(x, y). \quad (12.8)$$

Aquí, las funciones $\gamma_i: D \rightarrow \mathbb{R}$ (con $i = 1, 2$) son funciones continuas en el dominio D que es una región bidimensional del tipo 1. La condición $\gamma_1(x, y) = \gamma_2(x, y)$ implica que $(x, y) \in \partial D$, es decir, si las superficies $z = \gamma_1(x, y)$ y $z = \gamma_2(x, y)$ se intersectan, lo hacen sólo en puntos de la frontera.



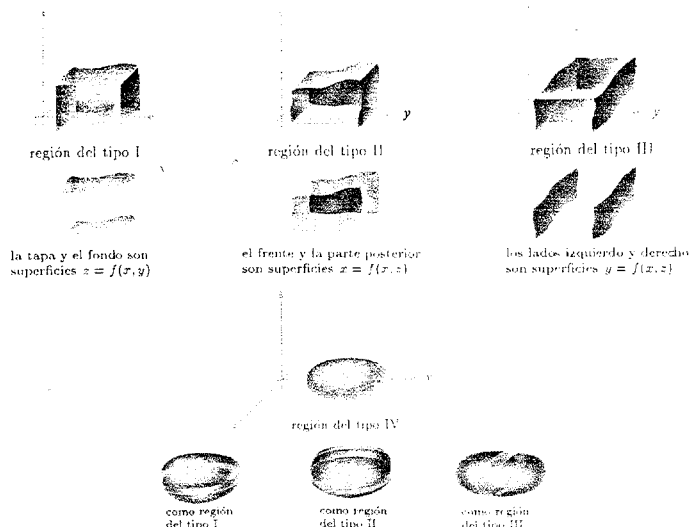
Una región tridimensional también se llamará del tipo I si puede expresarse como el conjunto de los puntos (x, y, z) tales que

$$c \leq y \leq d, \quad \Psi_1(y) \leq x \leq \Psi_2(y), \quad \gamma_1(x, y) \leq z \leq \gamma_2(x, y), \quad (12.9)$$

donde las funciones $\gamma_i: D \rightarrow \mathbb{R}$ (con $i = 1, 2$) son como antes y D es una región bidimensional del tipo 2.

Definición. Una región W es de **tipo II** si puede expresarse en la forma de la ecuación (12.8) o (12.9), intercambiando los papeles de x y z ; W es de **tipo III** si se puede expresar en la forma de la ecuación (12.8) o (12.9), con y y z intercambiados.

Finalmente una **región de tipo IV** es aquella que es del tipo I, tipo II y del tipo III al mismo tiempo. Un ejemplo de una región de tipo IV es la esfera de radio r , $x^2 + y^2 + z^2 \leq r^2$.



En la sección 12.3 demostramos los teoremas 12.5 y 12.6, para luego dar en la sección 12.4 la fórmula (12.6) válida para regiones bidimensionales de tipo 3. En el caso tridimensional se deberá llegar a fórmulas parecidas. Si W es del tipo I, tenemos

$$\begin{aligned} \int_W f(x, y, z) dV &= \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} \int_{\gamma_1(x, y)}^{\gamma_2(x, y)} f(x, y, z) dz dy dx = \\ &= \int_D \left[\int_{\gamma_1(x, y)}^{\gamma_2(x, y)} f(x, y, z) dz \right] dy dx, \end{aligned} \quad (12.10)$$

o

$$\begin{aligned} \int_W f(x, y, z) dV &= \int_c^d \int_{\Psi_1(y)}^{\Psi_2(y)} \int_{\gamma_1(x, y)}^{\gamma_2(x, y)} f(x, y, z) dz dx dy = \\ &= \int_D \left[\int_{\gamma_1(x, y)}^{\gamma_2(x, y)} f(x, y, z) dz \right] dx dy, \end{aligned} \quad (12.11)$$

según que W esté definida por la ecuación (12.8) o (12.9).

Si $f(x, y, z) = 1$ para todo $(x, y, z) \in W$, entonces obtenemos

$$\int_W f(x, y, z) dV = \int_W 1 dV = \text{Volumen } (W).$$

Si W es del tipo I y se puede aplicar la fórmula (12.10), obtenemos

$$\text{Volumen } (W) = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} \int_{\gamma_1(x, y)}^{\gamma_2(x, y)} dz dy dx = \int_a^b \int_{\Phi_1(x)}^{\Phi_2(x)} [\gamma_2(x, y) - \gamma_1(x, y)] dy dx.$$

12.5. LA INTEGRAL TRIPLE.

De la definición de integral como límite de sumas y los teoremas de límites, se pueden deducir algunas propiedades fundamentales de la integral triple, que son esencialmente las mismas que para la integral doble, vistas en la sección 12.2.

Propiedades de la integral triple.

Se supone que las funciones f y g son continuas en el paralelepípedo rectangular R .

1) Linealidad.

$$\int_R [f(x, y, z) + g(x, y, z)] dV = \int_R f(x, y, z) dV + \int_R g(x, y, z) dV;$$

2) Homogeneidad.

$$\int_R k f(x, y, z) dV = k \int_R f(x, y, z) dV, \quad \forall k \in \mathbb{R};$$

3) Monotonía. Si $f(x, y, z) \geq g(x, y, z)$ en R , entonces

$$\int_R f(x, y, z) dV \geq \int_R g(x, y, z) dV;$$

4) Aditividad. Si R_i , $i = 1, \dots, m$ son paralelepípedos rectangulares disjuntos entre sí, tales que f es integrable sobre cada R_i y si $Q = R_1 \cup R_2 \cup \dots \cup R_m$, entonces una función acotada $f: Q \rightarrow \mathbb{R}$ es integrable sobre Q y

$$\int_Q f(x, y, z) dV = \sum_{i=1}^m \int_{R_i} f(x, y, z) dV.$$

Teorema 12.10 Teorema del valor medio para integrales triples.

Sean f y g dos funciones continuas en una región $W \subset \mathbb{R}^3$. Si g es no negativa en W , existe un punto (x_0, y_0, z_0) en W para el cual

$$\int_W f(x, y, z) g(x, y, z) dx dy dz = f(x_0, y_0, z_0) \int_W g(x, y, z) dx dy dz.$$

Llameremos al valor $f(x_0, y_0, z_0)$ la **media ponderada por g de f en W** .

Si escogemos $g(x, y, z) = 1$ para todo $(x, y, z) \in W$, la fórmula anterior se reduce a

$$\int_W f(x, y, z) dx dy dz = f(x_0, y_0, z_0) \int_W dx dy dz = f(x_0, y_0, z_0) \cdot (\text{Volumen}(W)).$$

Diremos que $f(x_0, y_0, z_0)$ es el **valor medio de f en W** .

Ejemplo 1. Usa integrales triples para hallar el volumen del tetraedro T de vértices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$.

Solución:

Decidir cómo describir la región suele ser, con frecuencia, el paso más difícil en la evaluación de una integral triple. Una vez realizado esto de manera apropiada, queda sólo evaluar la integral triple dada usando una integral iterada equivalente.

El volumen T viene dado por la integral triple

$$\int_T dx dy dz$$

Para calcular esta integral, podemos proyectar T sobre cualquiera de los tres planos coordenados. Vamos a proyectar sobre el plano xy . La región de base es entonces el triángulo

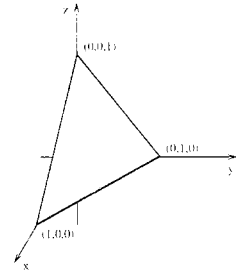
$$B = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}.$$

Dado que la cara superior es parte del plano inclinado $z = 1 - x - y$, tenemos que T es el conjunto de todas las ternas (x, y, z) tales que

$$T = \{(x, y, z) / 0 \leq x \leq 1, 0 \leq y \leq 1 - x, 0 \leq z \leq 1 - x - y\}.$$

Entonces,

$$\begin{aligned} V &= \int_T dx dy dz = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} dz dy dx = \int_0^1 \int_0^{1-x} (1-x-y) dy dx = \\ &= \int_0^1 \left[(1-x)y - \frac{y^2}{2} \right]_0^{1-x} dx = \frac{1}{2} \int_0^1 (1-x)^2 dx = -\frac{1}{6} [(1-x)^3]_0^1 = \frac{1}{6}. \end{aligned}$$



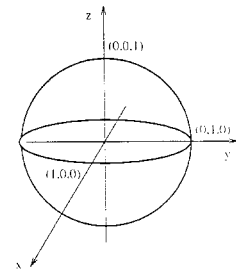
Ejemplo 2. Verifica la fórmula para el volumen de una bola $\int_S dx dy dz$, donde S es la esfera unitaria $x^2 + y^2 + z^2 \leq 1$.

Solución:

La región S es de tipo IV, y podemos entonces elegir cualquiera de las posibilidades para describirlas como región de tipo I, II o III. Si la suponemos como región de tipo I, podemos describir la esfera S como el conjunto de (x, y, z) que satisfacen

$$\begin{aligned} -1 \leq x \leq 1, \quad -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}, \\ -\sqrt{1-x^2-y^2} \leq z \leq \sqrt{1-x^2-y^2}. \end{aligned}$$

Aplicando la fórmula 12.10 y las propiedades de simetría, obtenemos



12.5. LA INTEGRAL TRIPLE.

$$\begin{aligned}
 V(S) &= \int_S dx dy dz = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} dz dy dx = \\
 &= 2 \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2} dy dx = 2 \int_{-1}^1 \pi \frac{1-x^2}{2} dx = \\
 &= \pi \int_{-1}^1 (1-x^2) dx = \pi \left[x - \frac{x^3}{3} \right]_{-1}^1 = \frac{4}{3} \pi.
 \end{aligned}$$

En el cálculo de la segunda integral en la segunda línea, se ha usado el siguiente resultado obtenido de una tabla de integrales:

$$\int \sqrt{a^2 - t^2} dt = \frac{t}{2} \sqrt{a^2 - t^2} + \frac{a^2}{2} \arcsen \frac{t}{a}, \quad a > 0,$$

con $a = \sqrt{1-x^2}$ y $t = y$.

Ejemplo 3. Sea W la región acotada por los planos $x = 0$, $y = 0$, $z = 2$ y la superficie $z = x^2 + y^2$, $x \geq 0$, $y \geq 0$. Calcula

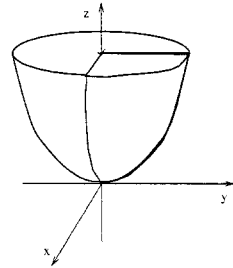
$$\int_W x dx dy dz.$$

Solución:

Si pensamos el sólido como región de tipo I, podemos describirlo como el conjunto de (x, y, z) que satisfacen

$$\begin{aligned}
 0 &\leq x \leq \sqrt{2}, \quad 0 \leq y \leq \sqrt{2-x^2}, \\
 x^2 + y^2 &\leq z \leq 2.
 \end{aligned}$$

Aplicando la fórmula 12.10, obtenemos



$$\begin{aligned}
 \int_W x dx dy dz &= \int_0^{\sqrt{2}} \int_0^{\sqrt{2-x^2}} \int_{x^2+y^2}^2 x dz dy dx = \int_0^{\sqrt{2}} \int_0^{\sqrt{2-x^2}} x [z]_{x^2+y^2}^2 dy dx = \\
 &= \int_0^{\sqrt{2}} \int_0^{\sqrt{2-x^2}} x (2 - x^2 - y^2) dy dx = \int_0^{\sqrt{2}} x \left[(2-x^2)y - \frac{y^3}{3} \right]_0^{\sqrt{2-x^2}} dx = \\
 &= \frac{2}{3} \int_0^{\sqrt{2}} x (2-x^2)^{3/2} dx = -\frac{1}{3} \int_0^{\sqrt{2}} (-2x) (2-x^2)^{3/2} dx = \\
 &= -\frac{2}{15} \left[(2-x^2)^{5/2} \right]_0^{\sqrt{2}} = \frac{2}{15} 2^{5/2} = \frac{8}{15} \sqrt{2}.
 \end{aligned}$$

También podemos pensar en W como una región de tipo II, es decir, podemos describirlo como el conjunto de (x, y, z) que satisfacen

$$0 \leq z \leq 2, \quad 0 \leq y \leq \sqrt{z}, \quad 0 \leq x \leq \sqrt{z - y^2}.$$

Aplicando la fórmula 12.11, obtenemos

$$\begin{aligned} \int_W x \, dx \, dy \, dz &= \int_0^2 \int_0^{\sqrt{z}} \int_0^{\sqrt{z-y^2}} x \, dx \, dy \, dz = \int_0^2 \int_0^{\sqrt{z}} \left[\frac{x^2}{2} \right]_0^{\sqrt{z-y^2}} dy \, dz = \\ &= \int_0^2 \int_0^{\sqrt{z}} \frac{z-y^2}{2} dy \, dz = \frac{1}{2} \int_0^2 \left[zy - \frac{y^3}{3} \right]_0^{\sqrt{z}} dz = \frac{1}{2} \int_0^2 \frac{2}{3} z^{3/2} dz = \\ &= \frac{2}{15} [z^{5/2}]_0^2 = \frac{2}{15} 2^{5/2} = \frac{8}{15} \sqrt{2}. \end{aligned}$$

Ejemplo 4. Integra $f(x, y, z) = yz$ sobre la parte del primer octante $x \geq 0, y \geq 0, z \geq 0$, determinada por el elipsoide

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1.$$

Solución:

Si suponemos el sólido T como región de tipo I, la frontera superior de T tiene por ecuación

$$z = \psi(x, y) = \frac{c}{ab} \sqrt{a^2 b^2 - b^2 x^2 - a^2 y^2}.$$

Esta superficie corta el plano xy a lo largo de la curva

$$y = \phi(x) = \frac{b}{a} \sqrt{a^2 - x^2}.$$

Podemos entonces considerar la región de base

$$D = \{(x, y) / 0 \leq x \leq a, 0 \leq y \leq \phi(x)\}.$$

Luego podemos caracterizar T como el conjunto de los (x, y, z) que satisfacen

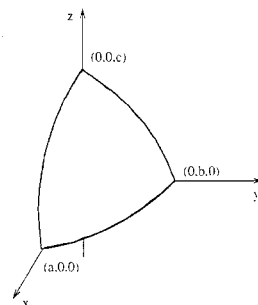
$$0 \leq x \leq a, \quad 0 \leq y \leq \phi(x) = \frac{b}{a} \sqrt{a^2 - x^2}, \quad 0 \leq z \leq \psi(x, y) = \frac{c}{ab} \sqrt{a^2 b^2 - b^2 x^2 - a^2 y^2}.$$

Aplicando la fórmula 12.10, obtenemos (comprobado evaluando la integral)

$$\int_T (yz) \, dx \, dy \, dz = \int_0^a \int_0^{\phi(x)} \int_0^{\psi(x,y)} (yz) \, dz \, dy \, dx = \frac{1}{15} a b^2 c^2.$$

También podemos pensar en T como una región de tipo II, es decir proyectando sobre el plano yz . En términos de y y z , la superficie curvada tiene por ecuación

$$x = \psi(y, z) = \frac{c}{ab} \sqrt{b^2 c^2 - c^2 y^2 - b^2 z^2}.$$



12.5. LA INTEGRAL TRIPLE.

Esta superficie corta el plano xy según la curva

$$y = \phi(z) = \frac{b}{c} \sqrt{c^2 - z^2}.$$

Podemos entonces considerar la región de base

$$D = \{(y, z) / 0 \leq z \leq c, 0 \leq y \leq \phi(z)\},$$

y caracterizar T como el conjunto de los (x, y, z) que satisfacen

$$0 \leq z \leq c, \quad 0 \leq y \leq \phi(z), \quad 0 \leq x \leq \psi(y, z).$$

Aplicando la fórmula 12.11, obtenemos (comprobado evaluando la integral)

$$\int_T (yz) \, dx \, dy \, dz = \int_0^c \int_0^{\phi(z)} \int_0^{\psi(y,z)} (yz) \, dx \, dy \, dz = \frac{1}{15} a b^2 c^2.$$

Ejemplo 5. Usa las integrales triples para hallar el volumen del sólido T limitado superiormente por el cilindro parabólico $z = 4 - y^2$ e inferiormente por el paraboloide elíptico $z = x^2 + 3y^2$.

Solución:

Resolviendo el sistema formado por las dos ecuaciones, tenemos

$$\begin{cases} z = 4 - y^2 \\ z = x^2 + 3y^2 \end{cases} \Rightarrow 4 - y^2 = x^2 + 3y^2 \Rightarrow$$

$$x^2 + 4y^2 = 4.$$

Esto significa que las dos superficies se cortan a lo largo de una curva en el espacio que está sobre el cilindro elíptico $x^2 + 4y^2 = 4$.

La proyección de esta curva sobre el plano xy es la elipse

$$x^2 + 4y^2 = 4.$$

Luego la región de base D será el conjunto de los puntos (x, y) tales que

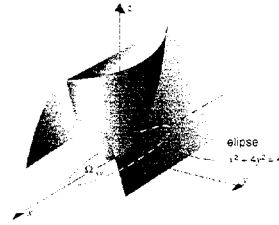
$$-2 \leq x \leq 2, \quad -\frac{1}{2} \sqrt{4 - x^2} \leq y \leq \frac{1}{2} \sqrt{4 - x^2}.$$

Luego podemos caracterizar T como el conjunto de los (x, y, z) que satisfacen

$$-2 \leq x \leq 2, \quad -\frac{1}{2} \sqrt{4 - x^2} \leq y \leq \frac{1}{2} \sqrt{4 - x^2}, \quad x^2 + 3y^2 \leq z \leq 4 - y^2.$$

Aplicando la fórmula 12.10 y teniendo en consideración la simetría del sólido, obtenemos (comprobado evaluando la integral)

$$V = \int_{-2}^2 \int_{-\sqrt{4-x^2}/2}^{\sqrt{4-x^2}/2} \int_{x^2+3y^2}^{4-y^2} dz \, dy \, dx = 4 \int_0^2 \int_0^{\sqrt{4-x^2}/2} \int_{x^2+3y^2}^{4-y^2} dz \, dy \, dx = 4\pi.$$



Ejercicios.

- 1) Dibuja la región de integración para cada integral, con su proyección en el plano xy . Expresa entonces cada integral triple como una o más integrales iteradas en las que la primera integración se efectúa respecto a y .

a)
$$\int_0^1 \int_0^{1-x} \int_0^{x+y} f(x, y, z) dz dy dx;$$

b)
$$\int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_0^{x+y} f(x, y, z) dz dy dx;$$

c)
$$\int_0^1 \int_0^1 \int_0^{x^2+y^2} f(x, y, z) dz dy dx.$$

- 2) Evalúa las siguientes integrales triple

a)
$$\int_W xyz dV, \quad W = [0, 1] \times [0, 1] \times [0, 1];$$

b)
$$\int_W xyz dV, \quad W = [0, a] \times [0, b] \times [0, c];$$

c)
$$\int_W x^2 dV, \quad W = [0, 1] \times [0, 1] \times [0, 1];$$

d)
$$\int_W e^{-xy} y dV, \quad W = [0, 1] \times [0, 1] \times [0, 1];$$

e)
$$\int_W (2x + 3y + z) dV, \quad W = [1, 2] \times [-1, 1] \times [0, 1];$$

f)
$$\int_W z e^{x+y} dV, \quad W = [0, 1] \times [0, 1] \times [0, 1];$$

g)
$$\int_W x^2 \cos z dV, \quad W \text{ es la región acotada por los planos } z = 0, z = \pi, y = 0, y = 1, x = 0 \text{ y } x + y = 1;$$

h)
$$\int_0^a \int_0^b \int_0^c dx dy dz;$$

i)
$$\int_0^1 \int_0^x \int_0^y y dz dy dx;$$

j)
$$\int_0^1 \int_1^{2y} \int_0^x (x + 2z) dz dx dy;$$

k)
$$\int_0^1 \int_{1-x}^{1-x} \int_0^{xy} 4z dz dy dx;$$

l)
$$\int_0^1 \int_0^{2x} \int_{x^2+y^2}^{x+y} dz dy dx.$$

- 3) Halla el volumen de la región acotada por $z = x^2 + 3y^2$ y $z = 9 - x^2$.

- 4) Cambia el orden de integración en la integral $\int_0^1 \int_0^x \int_0^y f(x, y, z) dz dy dx$ para obtener otras cinco formas de la respuesta. Esboza la región de integración.

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

5) Sea T un sólido cuyo volumen es igual a $V = \iiint_T dx dy dz$. Dibuja T y rellena los blancos en las siguientes integrales

$$\begin{aligned} \text{a)} \quad V &= \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} dy dx dz; & \text{b)} \quad V &= \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} dy dz dx; \\ \text{c)} \quad V &= \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} dx dy dz + \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} \int_{\diamond}^{\diamond} dx dy dz; \end{aligned}$$

si

$$\begin{aligned} \text{i)} \quad V &= \int_0^2 \int_0^{9-x^2} \int_0^{2-x} dz dy dx; & \text{ii)} \quad V &= \int_0^3 \int_x^{6-x} \int_0^{2x} dz dy dx. \end{aligned}$$

12.6 Cambio de variables en la integral múltiple.

Antes de estudiar la fórmula del cambio de variables para integrales múltiples, es necesario ocuparnos del efecto de las funciones de \mathbb{R}^2 a \mathbb{R}^2 sobre los subconjuntos de \mathbb{R}^2 . En otras palabras, primero debemos estudiar:

12.6.1 La geometría de las funciones de \mathbb{R}^2 a \mathbb{R}^2 .

Sea D^* un subconjunto de \mathbb{R}^2 y sea $T: D^* \rightarrow \mathbb{R}^2$ una función continuamente diferenciable, que lleva puntos de D^* a puntos en \mathbb{R}^2 . Denotemos este conjunto imagen por D , entonces

$$D = T(D^*) = \{(x, y) \in \mathbb{R}^2 / (x, y) = T(x^*, y^*) \text{ para algún } (x^*, y^*) \in D^*\}.$$

Una manera de entender la geometría de la función T es ver cómo deforma o cambia el conjunto original D^* .

Ejemplo 1. Sea $D^* \subset \mathbb{R}^2$ el rectángulo $D^* = [0, 1] \times [0, 2\pi]$. Halla el conjunto imagen $D = T(D^*)$, si T viene definida por $T(r, \theta) = (r \cos \theta, r \sin \theta)$.

Solución:

El dominio D^* resulta ser el conjunto $\{(r, \theta) \in \mathbb{R}^2 / 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}$, que es un rectángulo en el plano $r\theta$.

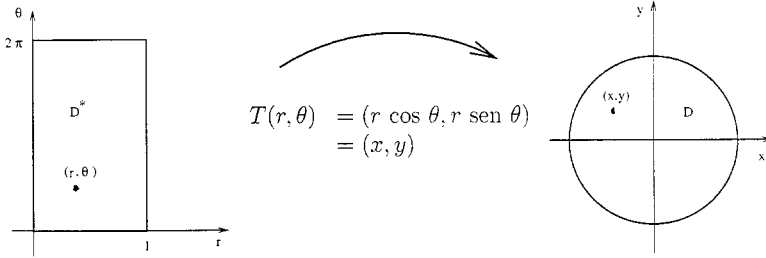
Los puntos (x, y) del conjunto imagen $D = T(D^*) \subset \mathbb{R}^2$ vienen definidos por $(x, y) = (r \cos \theta, r \sin \theta)$. Dado que de

$$x^2 + y^2 = r^2 \cos^2 \theta + r^2 \sin^2 \theta = r^2 \leq 1$$

se sigue que todos los puntos de D están contenidos en el disco unitario. Además, cualquier punto (x, y) en el disco unitario se puede escribir como $(r \cos \theta, r \sin \theta)$ para algunos valores $0 \leq r \leq 1$ y $0 \leq \theta \leq 2\pi$. Por todo esto, D es el disco unitario.

Es interesante notar que todo valor con $r = 0$ tiene como imagen el punto $(0, 0)$ en el plano xy .

Esta función T nos da un **cambio de variables a coordenadas polares**, ver la fórmula (9.1) dada en la sección 9.4.



Ejemplo 2. Sea $D^* \subset \mathbb{R}^2$ el cuadrado $D^* = [-1, 1] \times [-1, 1]$. Halla el conjunto imagen $D = T(D^*)$, si T viene definida por $T(u, v) = \left(\frac{u+v}{2}, \frac{u-v}{2} \right)$.

Solución:

Determinemos primero el efecto de T en los cuatro lados del cuadrado. Empezamos parametrizando los segmentos

$$\begin{aligned} \vec{\sigma}_1(t) &= (t, 1) & -1 \leq t \leq 1, & & \vec{\sigma}_2(t) &= (1, t) & -1 \leq t \leq 1, \\ \vec{\sigma}_3(t) &= (t, -1) & -1 \leq t \leq 1, & & \vec{\sigma}_4(t) &= (-1, t) & -1 \leq t \leq 1. \end{aligned}$$

Luego

$$T(\vec{\sigma}_1(t)) = \left(\frac{t+1}{2}, \frac{t-1}{2} \right) \quad \text{con} \quad -1 \leq t \leq 1$$

es una parametrización de la recta

$$y = x - 1 \quad \text{con} \quad 0 \leq x \leq 1.$$

Para verlo eliminamos el parámetro t del sistema

$$\begin{cases} x = \frac{t+1}{2} \\ y = \frac{t-1}{2} \end{cases} \Leftrightarrow \begin{cases} t = 2x - 1 \\ t = 2y + 1 \end{cases} \Leftrightarrow 2x - 1 = 2y + 1 \Leftrightarrow y = x - 1;$$

además, mientras t varía desde -1 hasta 1 , la variable x varía desde 0 hasta 1 .

De manera análoga, los otros tres lados del cuadrado dan como resultado que:

$$T(\vec{\sigma}_2(t)) \text{ es una parametrización de la recta } y = 1 - x \quad \text{con} \quad 0 \leq x \leq 1,$$

$$T(\vec{\sigma}_3(t)) \text{ es una parametrización de la recta } y = x + 1 \quad \text{con} \quad -1 \leq x \leq 0.$$

$$T(\vec{\sigma}_4(t)) \text{ es una parametrización de la recta } y = -x - 1 \quad \text{con} \quad -1 \leq x \leq 0.$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Parece entonces razonable pensar que la aplicación T va a “girar” el cuadrado D para colocarlo sobre el cuadrado D cuyos vértices son $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$.

Para probar que así sucede, sea $-1 \leq \alpha \leq 1$ y sea L_α una recta fija parametrizada por $\vec{\sigma}(t) = (\alpha, t)$, con $-1 \leq t \leq 1$; entonces

$$T(\vec{\sigma}(t)) = \left(\frac{\alpha + t}{2}, \frac{\alpha - t}{2} \right)$$

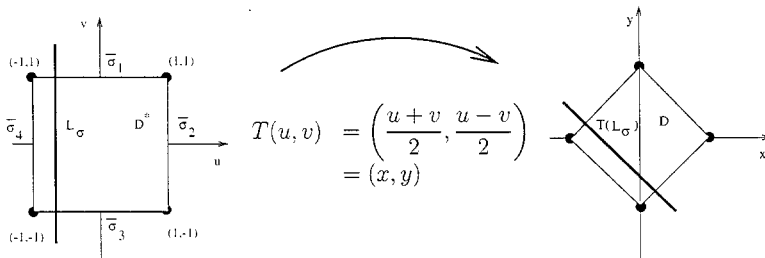
es una parametrización de la recta

$$y = -x + \alpha \quad \text{con} \quad \frac{\alpha - 1}{2} \leq x \leq \frac{\alpha + 1}{2}.$$

Esta recta comienza, para $t = -1$, y termina, para $t = 1$, respectivamente, en los puntos

$$\left(\frac{\alpha - 1}{2}, \frac{\alpha + 1}{2} \right) \quad \left(\frac{\alpha + 1}{2}, \frac{\alpha - 1}{2} \right)$$

que son puntos que están sobre las rectas $T \circ \vec{\sigma}_3$ y $T \circ \vec{\sigma}_1$, respectivamente. Así, cuando α varía entre -1 y 1 , la recta L_α barre el cuadrado D^* mientras que $T(L_\alpha)$ barre el cuadrado D determinado por los vértices $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$.



El teorema siguiente proporciona una manera fácil de describir la imagen $T(D^*)$.

Teorema 12.11 Sea A una matriz 2×2 con $\det A \neq 0$ y T una transformación lineal de \mathbb{R}^2 a \mathbb{R}^2 dada por $T(\vec{x}) = A\vec{x}$. Entonces, T transforma paralelogramos en paralelogramos y vértices en vértices. Más aún, si $T(D^*)$ es un paralelogramo, D^* debe ser un paralelogramo.

Definición. La función T es uno a uno (o inyectiva) en D^* si para (u, v) y (u', v') en D^* , $T(u, v) = T(u', v')$ implica que $u = u'$ y $v = v'$.

Geoméricamente, esto significa que dos puntos diferentes de D^* no van a ser el mismo punto de D bajo la aplicación T .

Ejemplo 1. La aplicación $T(x, y) = (x^2 + y^2, y^4)$ no es uno a uno porque $T(1, -1) = (2, 1) = T(1, 1)$, aunque $(1, -1) \neq (1, 1)$.

Ejemplo 2. Volvamos a considerar la aplicación que nos da el cambio de variables a coordenadas polares, $T(r, \theta) = (r \cos \theta, r \sin \theta)$. Demuestra que T no es uno a uno si su dominio es todo \mathbb{R}^2 .

Solución:

Si $\theta_1 \neq \theta_2$, entonces $T(0, \theta_1) = (0, 0) = T(0, \theta_2)$, de modo que T no puede ser uno a uno. Esta observación implica que si L es el lado del rectángulo $D^* = [0, 1] \times [0, 2\pi]$ donde $r = 0$ y $0 \leq \theta \leq 2\pi$, entonces T manda todo el lado L a un solo punto, el centro del disco unitario D . Sin embargo si consideramos el conjunto $S^* = (0, 1] \times [0, 2\pi)$, entonces $T: S^* \rightarrow S$ es uno a uno.

Ejemplo 3. Volvamos a considerar la aplicación $T(u, v) = \left(\frac{u+v}{2}, \frac{u-v}{2}\right)$, que vimos en un ejemplo anterior como aplicación de $D^* \subset \mathbb{R}^2$ en \mathbb{R}^2 , donde D^* es el cuadrado $[-1, 1] \times [-1, 1]$. Muestra que como aplicación $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es uno a uno.

Solución:

Supongamos que $T(u, v) = T(u', v')$; entonces

$$\left(\frac{u+v}{2}, \frac{u-v}{2}\right) = \left(\frac{u'+v'}{2}, \frac{u'-v'}{2}\right) \Rightarrow \begin{cases} u+v = u'+v' \\ u-v = u'-v' \end{cases} \Rightarrow \begin{cases} u = u' \\ v = v' \end{cases}$$

Esto muestra que T es uno a uno con dominio todo \mathbb{R}^2 . En realidad, como T es lineal y $T(\vec{x}) = A\vec{x}$, donde A es una matriz 2×2 , bastaría demostrar que $\det A \neq 0$ (se propone como ejercicio).

Definición. La función T es **sobreyectiva** en D si para cada punto $(x, y) \in D$ existe al menos un punto (u, v) en el dominio D^* , tal que $T(u, v) = (x, y)$.

Así, si T es sobreyectiva, podemos resolver la ecuación $T(u, v) = (x, y)$ para (u, v) dado que $(x, y) \in D$. Si T es, además, uno a uno, esta solución es única.

En los ejemplos anteriores hemos determinado la imagen $D = T(D^*)$ de una región D^* bajo una función T . Lo que nos interesará en lo que sigue será, en parte, el problema inverso: dada D y una función uno a uno T de \mathbb{R}^2 a \mathbb{R}^2 , hallar D^* tal que $T(D^*) = D$.

Si tenemos una región D dada y una función T , la determinación de una región D^* tal que $T(D^*) = D$ será posible sólo cuando para todo $(x, y) \in D$ exista un (u, v) en el dominio de T tal que $T(u, v) = (x, y)$, esto es T debe ser sobreyectiva. Esto no siempre es posible.

Ejemplo 1. Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(u, v) = (u, 0)$. Sea $D = [0, 1] \times [0, 1]$. Como T lleva todo \mathbb{R}^2 a uno de los ejes, es imposible hallar D^* tal que $T(D^*) = D$.

Ejemplo 2. Volvamos a considerar la aplicación $T(u, v) = \left(\frac{u+v}{2}, \frac{u-v}{2}\right)$. Sea D el cuadrado cuyos vértices son $(1, 0)$, $(0, 1)$, $(-1, 0)$ y $(0, -1)$. Halla D^* tal que $T(D^*) = D$.

Solución:

Como T es lineal y $T(\vec{x}) = A\vec{x}$, donde A es una matriz 2×2 que satisface $\det A \neq 0$, entonces $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es sobreyectiva (demostradlo) y así, es posible hallar D^* . Ahora por el teorema 12.11, D^* debe ser un paralelogramo, luego para hallar D^* basta hallar

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

los cuatro puntos que van a dar los vértices de D ; entonces al conectar estos puntos, habremos hallado D^* .

Para el vértice $(1, 0)$ de D , debemos resolver

$$T(u, v) = \left(\frac{u+v}{2}, \frac{u-v}{2} \right) = (1, 0) \Rightarrow \begin{cases} u+v=2 \\ u-v=0 \end{cases} \Rightarrow \begin{cases} u=1 \\ v=1. \end{cases}$$

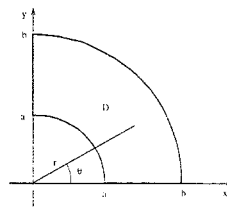
Así, $(u, v) = (1, 1)$ es un vértice de D^* . Resolviendo para los otros vértices, hallamos que $D^* = [-1, 1] \times [-1, 1]$.

Ejemplo 3. Sea D la región que está en el primer cuadrante comprendida entre los arcos de las circunferencias $x^2 + y^2 = a^2$ y $x^2 + y^2 = b^2$, con $0 < a < b$. Sea T la transformación a coordenadas polares dada por $T(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$. Halla D^* tal que $T(D^*) = D$.

Solución:

Los contornos dados tienen, en coordenadas polares, ecuaciones $r = a$ y $r = b$, respectivamente. La región D es tal que $a^2 \leq x^2 + y^2 \leq b^2$, y como $x^2 + y^2 = r^2$, se sigue que $a \leq r \leq b$. Claramente, para esta región el ángulo θ varía entre 0 y $\pi/2$.

Así, es $D^* = [a, b] \times \left[0, \frac{\pi}{2}\right]$. Además, T es uno a uno.



12.6.2 Cambio de variables en la integral doble.

Pasemos ahora a estudiar el siguiente problema: dadas dos regiones D y D^* del tipo 1 ó 2 en \mathbb{R}^2 , dada una función diferenciable T en D^* con imagen D , esto es $T(D^*) = D$ y dada cualquier otra función integrable con valores reales, $f: D \rightarrow \mathbb{R}$, quisiéramos expresar $\int_D f(x, y) dA$ como una integral sobre D^* de la función compuesta $f \circ T$.

Supongamos que D^* y D son dos subconjuntos de \mathbb{R}^2 , ambos del tipo 1, aunque pensados el primero como subconjunto del plano uv y el segundo del plano xy . Entonces, la función T está dada por dos funciones coordenadas:

$$T(u, v) = (x(u, v), y(u, v)) \quad \text{para } (u, v) \in D^*.$$

Como primera conjetura podríamos poner

$$\int_D f(x, y) dx dy \stackrel{?}{=} \int_{D^*} f(x(u, v), y(u, v)) du dv$$

donde $f \circ T(u, v) = f(x(u, v), y(u, v))$ es la función compuesta definida en D^* . Sin embargo, si consideramos la función $f: D \rightarrow \mathbb{R}$, dada por $f(x, y) = 1$, entonces esto nos daría

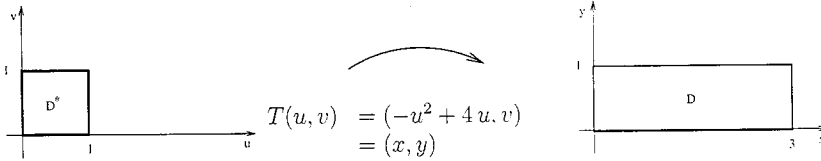
$$A(D) = \int_D dx dy \stackrel{?}{=} \int_{D^*} du dv = A(D^*),$$

esto es la igualdad de las áreas de los dominios D y D^* . Es fácil ver que esto se cumpliría sólo para algunos casos particulares y no para una función genérica T .

Ejemplo. Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(u, v) = (-u^2 + 4u, v)$. Restringe T al cuadrado unitario $D^* = [0, 1] \times [0, 1]$, y comprueba que $A(D) \neq A(D^*)$.

Solución:

Si procedemos de la misma manera que en los ejemplos anteriores, tendremos que la imagen mediante T del dominio D^* es el rectángulo $D = T(D^*) = [0, 3] \times [0, 1]$, y claramente $A(D) = 1 \neq A(D^*) = 3$.



Lo que se necesita es una medida de cómo la transformación $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ distorsiona el área de la región. Esto está dado por el determinante jacobiano que se define como sigue.

Definición. Sea $T: D^* \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ una transformación C^1 dada por $x = x(u, v)$, $y = y(u, v)$. El **jacobiano** de T es el determinante de la matriz derivada $DT(x, y)$ de T :

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}.$$

Si hacemos las restricciones adecuadas a la función T , podemos mostrar que el área de $D = T(D^*)$ se obtiene integrando el valor absoluto del jacobiano sobre D^* ; esto es,

$$A(D) = \int_D dx dy = \int_{D^*} \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv. \quad (12.12)$$

Ejemplo. La función $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que nos da el cambio a coordenadas polares viene dada por $T(r, \theta) = (r \cos \theta, r \sin \theta)$. Halla el área de $D = T(D^*)$ si $D^* = [0, 1] \times [0, 2\pi]$.

Solución:

Según la fórmula 12.12 necesitamos el jacobiano de la transformación T . Dado que la función que transforma coordenadas polares en coordenadas rectangulares, y viceversa, viene dada por

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

su jacobiano será (notar que aquí r y θ juegan el papel de u y v , respectivamente)

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r (\cos^2 \theta + \sin^2 \theta) = r.$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Luego, usando la fórmula 12.12 tenemos

$$\begin{aligned} A(D) &= \int_D dx dy = \int_{D^*} \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| dr d\theta = \int_{D^*} r dr d\theta = \\ &= \int_0^{2\pi} \int_0^1 r dr d\theta = \int_0^{2\pi} \left[\frac{r^2}{2} \right]_0^1 d\theta = \frac{1}{2} \int_0^{2\pi} d\theta = \frac{1}{2} (2\pi) = \pi. \end{aligned}$$

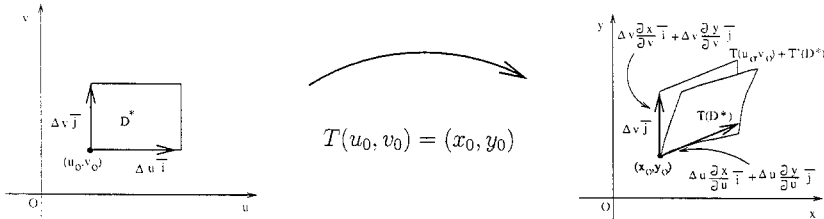
Este resultado concuerda con la información obtenida anteriormente: en el primer ejemplo de esta sección vimos que $D = T(D^*)$ es el disco unitario dado por $x^2 + y^2 \leq 1$, cuya área es claramente π (área del disco = πr^2 , con $r = 1$).

No es fácil probar de manera rigurosa la fórmula 12.12, es decir, demostrar que el determinante jacobiano es una medida de cómo distorsiona el área una transformación. Sin embargo, visto de manera adecuada, es plausible. Recordemos que $A(D) = \int_D dx dy$ se obtuvo al dividir D en rectángulos pequeños, sumando sus áreas y tomando el límite de esta suma conforme el tamaño de los rectángulos tiende a cero. El problema es que T transforma rectángulos en regiones cuyas áreas no son fáciles de calcular. Una herramienta útil es la derivada de T , que como sabemos de la sección 10.3 es la mejor aproximación lineal a la función T .

Consideremos, entonces, un pequeño rectángulo D^* en el plano uv . Denotemos por T' la derivada de T evaluada en el punto (u_0, v_0) , de modo que T' es una matriz 2×2 . Por lo visto en la sección 10.3 (ver la definición de diferenciabilidad), sabemos que una buena aproximación a $T(u, v)$ está dada por

$$T(u, v) \approx T(u_0, v_0) + \begin{pmatrix} \frac{\partial T_1}{\partial u} & \frac{\partial T_1}{\partial v} \\ \frac{\partial T_2}{\partial u} & \frac{\partial T_2}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} = T(u_0, v_0) + \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix},$$

donde $\Delta u = u - u_0$ y $\Delta v = v - v_0$.



Pero esta correspondencia manda D^* a un paralelogramo con vértice en $T(u_0, v_0)$ y con lados adyacentes dados por los vectores

$$T'(\Delta u \vec{i}) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} \Delta u \\ 0 \end{pmatrix} = \Delta u \cdot \begin{pmatrix} \frac{\partial x}{\partial u} \\ \frac{\partial y}{\partial u} \end{pmatrix} = \Delta u \vec{T}_u,$$

$$T'(\Delta v \vec{j}) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \Delta v \end{pmatrix} = \Delta v \cdot \begin{pmatrix} \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial v} \end{pmatrix} = \Delta v \vec{T}_v,$$

donde los vectores

$$\vec{T}_u = \frac{\partial x}{\partial u} \vec{i} + \frac{\partial y}{\partial u} \vec{j} \quad \text{y} \quad \vec{T}_v = \frac{\partial x}{\partial v} \vec{i} + \frac{\partial y}{\partial v} \vec{j}$$

están evaluados en (u_0, v_0) .

Sabemos de la sección 9.3 que el área del paralelogramo con lados iguales a los vectores $\vec{v}_1 = a\vec{i} + b\vec{j}$ y $\vec{v}_2 = c\vec{i} + d\vec{j}$ es igual al valor absoluto del determinante

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix}.$$

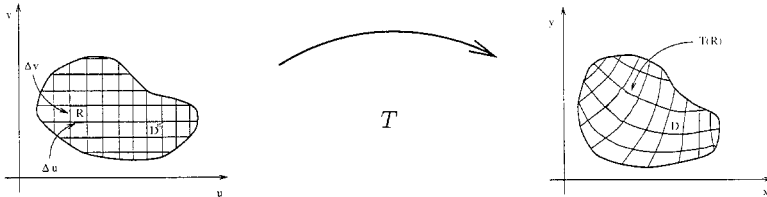
Así, el área $T(D^*)$ es aproximadamente igual al valor absoluto de

$$\begin{vmatrix} \frac{\partial x}{\partial u} \Delta u & \frac{\partial x}{\partial v} \Delta v \\ \frac{\partial y}{\partial u} \Delta u & \frac{\partial y}{\partial v} \Delta v \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \Delta u \Delta v = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \Delta u \Delta v$$

evaluado en (u_0, v_0) .

Si ahora consideremos una partición de D^* en rectángulos pequeños con lados de longitudes Δu y Δv , las imágenes de estos rectángulos están aproximadas por los paralelogramos con lados $\vec{T}_u \Delta u$ y $\vec{T}_v \Delta v$, y por lo tanto con área $\left| \frac{\partial(x, y)}{\partial(u, v)} \right| \Delta u \Delta v$. Sumando sobre todos

los paralelogramos y pasando al límite, la suma se convierte en $\int_{D^*} \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$.



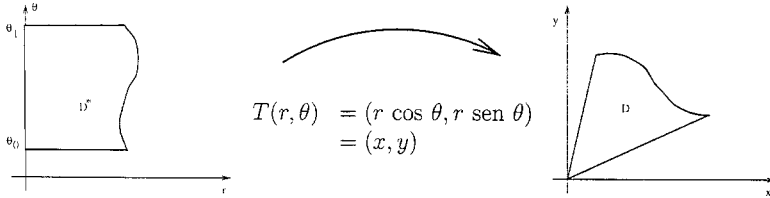
Ejemplo. Volvamos a considerar la función $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que nos da el cambio a coordenadas polares: $T(r, \theta) = (r \cos \theta, r \sin \theta)$. Sea D la región elemental en el plano xy acotada por la gráfica de una ecuación polar $r = f(\theta)$, donde $\theta_0 \leq \theta \leq \theta_1$ y $f(\theta) \geq 0$. En el plano $r\theta$ consideremos la región D^* del tipo 2 donde $\theta_0 \leq \theta \leq \theta_1$ y $0 \leq r \leq f(\theta)$. Bajo la transformación T , la región D^* va a dar sobre D . Usar la fórmula (12.12) para calcular el área de D .

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Solución:

$$\begin{aligned} A(D) &= \int_D dx dy = \int_{D^*} \left| \frac{\partial(x,y)}{\partial(r,\theta)} \right| dr d\theta = \int_{D^*} r dr d\theta = \\ &= \int_{\theta_0}^{\theta_1} \left[\int_0^{f(\theta)} r dr \right] d\theta = \int_{\theta_0}^{\theta_1} \left[\frac{r^2}{2} \right]_0^{f(\theta)} d\theta = \int_{\theta_0}^{\theta_1} \frac{f(\theta)^2}{2} d\theta, \end{aligned}$$

que es la misma fórmula que vimos en la sección 3.6.



En el teorema 3.4 se dió la fórmula (3.2) para el cambio de variable en el caso de una función de una variable

$$\int_a^b f(x(u)) x'(u) du = \int_{x(a)}^{x(b)} f(x) dx \quad (12.13)$$

válida si f es continua y la aplicación que transforma u en $x(u)$ es continuamente diferenciable en $[a, b]$. Queremos ahora reescribir esta última fórmula para luego generalizar el resultado a las integrales dobles.

Supongamos que la función de clase C^1 que transforma u en $x(u)$ es uno a uno (o inyectiva) en $[a, b]$. Tenemos

$$\frac{dx}{du} \geq 0 \quad \text{ó} \quad \frac{dx}{du} \leq 0.$$

(Nótese que si, por ejemplo, la derivada es primero positiva y luego negativa, la función $x = x(u)$ crece y después decrece, y así, no es uno a uno).

Denotemos por I^* el intervalo $[a, b]$ y por I el intervalo cerrado con extremos $x(a)$ y $x(b)$ (es decir, $I = [x(a), x(b)]$ si la función $x(u)$ es creciente y $I = [x(b), x(a)]$ si la función $x(u)$ es decreciente). Entonces podemos reescribir la fórmula (12.13) como

$$\int_{I^*} f(x(u)) \left| \frac{dx}{du} \right| du = \int_I f(x) dx.$$

Ésta es la fórmula que se generaliza para integrales dobles: el intervalo I^* se convierte en la región D^* , el intervalo I se convierte en D y el valor absoluto de la derivada se sustituye por el determinante del jacobiano. El enunciado formal de este resultado (sin demostración) es el siguiente:

Teorema 12.12 *Cambio de variables para integrales dobles.*

Sean D y D^* regiones elementales en el plano y sea $T: D^* \rightarrow D$ de clase C^1 y además es uno a uno en D^* . Más aún, sea $D = T(D^*)$. Entonces, para cualquier función integrable $f: D \rightarrow \mathbb{R}$, tenemos

$$\int_D f(x, y) dx dy = \int_{D^*} f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv. \quad (12.14)$$

Uno de los propósitos de este teorema de cambio de variables es proporcionar un método mediante el cual se puedan simplificar algunas integrales dobles. A veces puede ser muy difícil hallar una integral doble debido a que es complicado el integrando f o la región D ; escogiendo una aplicación T adecuada puede ser fácil evaluar la nueva integral con el nuevo integrando $f \circ T$ en la nueva región D^* . Desafortunadamente, si no se escoge bien y con cuidado la función T , el problema puede complicarse más aún.

Ejemplo 1. Sea P el paralelogramo acotado por $y = 2x$, $y = 2x - 2$, $y = x$ e $y = x + 1$. Evalúa

$$\int_P xy dx dy$$

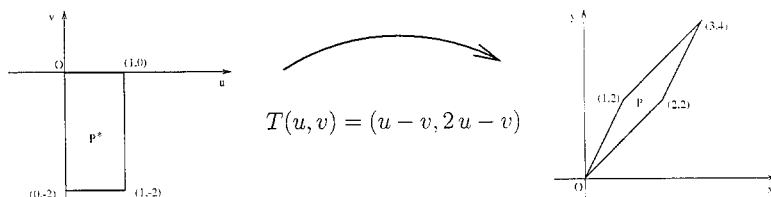
usando $T(u, v) = (u - v, 2u - v)$, en otras palabras haciendo el cambio de variables

$$x = u - v, \quad y = 2u - v.$$

Solución:

La aplicación T es uno a uno (comprobadlo) y está diseñada de modo que convierte el rectángulo P^* acotado por $v = 0$, $v = -2$, $u = 0$ y $u = 1$ en P . Al usar T se simplifica la región de integración de P a P^* . Más aún,

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 2 & -1 \end{vmatrix} = -1 + 2 = 1.$$



Por lo tanto

$$\begin{aligned} \int_P xy dx dy &= \int_{P^*} (u - v)(2u - v) |1| du dv = \int_{-2}^0 \int_0^1 (2u^2 - 3uv + v^2) du dv = \\ &= \int_{-2}^0 \left[\frac{2}{3} u^3 - 3 \frac{u^2}{2} v + uv^2 \right]_0^1 dv = \int_{-2}^0 \left[\frac{2}{3} - \frac{3}{2} v + v^2 \right] dv = \left[\frac{2}{3} v - \frac{3}{4} v^2 + \frac{v^3}{3} \right]_{-2}^0 = 7. \end{aligned}$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Ejemplo 2. Calcula

$$\int_R (x + y) dx dy$$

donde R es el paralelogramo delimitado por las rectas

$$x + y = 0, \quad x + y = 1, \quad 2x - y = 0, \quad 2x - y = 3.$$

Solución:

Las fronteras del paralelogramo sugieren que se haga

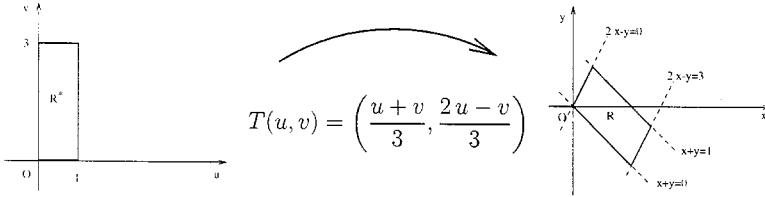
$$u = x + y, \quad v = 2x - y.$$

Queremos expresar x e y en función de u y v . Dado que

$$\begin{cases} u + v = (x + y) + (2x - y) = 3x \\ 2u - v = (2x + 2y) - (2x - y) = 3y \end{cases} \Rightarrow x = \frac{u + v}{3}, \quad y = \frac{2u - v}{3}.$$

Esta transformación aplica el rectángulo $R^* = [0, 1] \times [0, 3]$ en el paralelogramo R original. Es decir obtenemos la transformación

$$T(u, v) = \left(\frac{u + v}{3}, \frac{2u - v}{3} \right) = (x, y).$$



El jacobiano de esta transformación es

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{vmatrix} = -\frac{1}{3}.$$

Por lo tanto

$$\begin{aligned} \int_R (x + y)^2 dx dy &= \int_{R^*} u^2 \left| -\frac{1}{3} \right| du dv = \frac{1}{3} \int_0^3 \int_0^1 u^2 du dv = \\ &= \frac{1}{3} \left(\int_0^3 dv \right) \left(\int_0^1 u^2 du \right) = \frac{1}{3} [v]_0^3 \left[\frac{u^3}{3} \right]_0^1 = \frac{1}{3}. \end{aligned}$$

Ejemplo 3. Calcula

$$\int_R xy \, dx \, dy$$

donde R es la región del primer cuadrante delimitada por las curvas

$$x^2 + y^2 = 4, \quad x^2 + y^2 = 9, \quad x^2 - y^2 = 1, \quad x^2 - y^2 = 4.$$

Solución:

Las fronteras de la región considerada sugieren que se haga

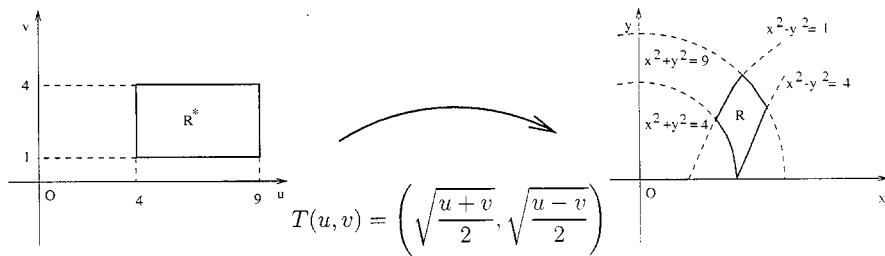
$$u = x^2 + y^2, \quad v = x^2 - y^2.$$

Como en el ejemplo anterior queremos expresar x e y en función de u y v . Dado que

$$\begin{cases} u + v = (x^2 + y^2) + (x^2 - y^2) = 2x^2 \\ u - v = (x^2 + y^2) - (x^2 - y^2) = 2y^2 \end{cases} \Rightarrow x = \sqrt{\frac{u+v}{2}}, \quad y = \sqrt{\frac{u-v}{2}}.$$

Esta transformación aplica el rectángulo $R^* = [4, 9] \times [1, 4]$ en la región R original. Es decir obtenemos la transformación

$$T(u, v) = \left(\sqrt{\frac{u+v}{2}}, \sqrt{\frac{u-v}{2}} \right) = (x, y).$$



El jacobiano de esta transformación es (comprobadlo)

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = -\frac{1}{4\sqrt{u^2 - v^2}}.$$

Por lo tanto

$$\begin{aligned} \int_R xy \, dx \, dy &= \int_{R^*} \left(\sqrt{\frac{u+v}{2}} \right) \left(\sqrt{\frac{u-v}{2}} \right) \left| -\frac{1}{4\sqrt{u^2 - v^2}} \right| du \, dv = \int_1^4 \int_4^9 \frac{1}{8} du \, dv = \\ &= \frac{1}{8} (\text{área de } R^*) = \frac{1}{8} \left(\int_1^4 dv \right) \left(\int_4^9 du \right) = \frac{1}{8} (3) (5) = \frac{15}{8}. \end{aligned}$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Ejemplo 4. Evalúa

$$\int_D \log(x^2 + y^2) dx dy$$

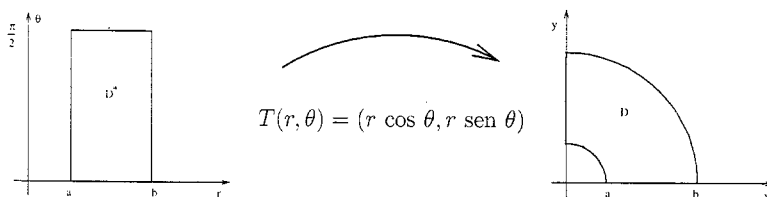
donde D es la región en el primer cuadrante que está entre los arcos de circunferencias

$$x^2 + y^2 = a^2 \quad x^2 + y^2 = b^2 \quad \text{con } 0 < a < b.$$

Solución:

Estas circunferencias tienen las ecuaciones sencillas $r = a$ y $r = b$ en coordenadas polares. Además, $r^2 = x^2 + y^2$ aparece en el integrando, de manera que un cambio a coordenadas polares simplificará tanto el integrando como la región de integración.

Ya sabemos que la transformación a coordenadas polares manda el rectángulo D^* , dado por $a \leq r \leq b$, $0 \leq \theta \leq \pi/2$ sobre la región D .



Sabiendo que esta transformación es uno a uno en D^* , el teorema 12.12 nos permite escribir, sabiendo que

$$(x, y) = T(r, \theta) = (r \cos \theta, r \sin \theta) \Rightarrow \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| = r,$$

$$\begin{aligned} \int_D \log(x^2 + y^2) dx dy &= \int_{D^*} \log r^2 \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| dr d\theta = \int_a^b \int_0^{\pi/2} r \log r^2 d\theta dr = \\ &= \frac{\pi}{2} \int_a^b r \log r^2 dr = \frac{\pi}{2} \int_a^b 2r \log r dr = \frac{\pi}{2} \left[b^2 \log b - a^2 \log a - \frac{1}{2}(b^2 - a^2) \right], \end{aligned}$$

donde se ha usado la fórmula

$$\int x \log x dx = \frac{x^2}{2} \log x - \frac{x^2}{4}$$

obtenida por el método de integración por partes.

Nótese que, como ya sabemos, la transformación de coordenadas cartesianas a coordenadas polares no es uno a uno si consideremos el rectángulo D^* definido por $0 \leq \theta \leq 2\pi$, $0 \leq r \leq a$ porque transforma todos los puntos con $r = 0$ en el único punto $(0, 0)$. Sin embargo el teorema del cambio de variables se cumple en este caso. Básicamente la razón es que el conjunto de puntos donde T no es uno a uno está en un lado de D^* , que es la gráfica de una curva suave y por lo tanto, para los propósitos de integración, puede despreciarse.

Veamos ahora un ejemplo de uso de coordenadas polares, más complicado.

Ejemplo. Evalúa

$$\int_R \sqrt{x^2 + y^2} dx dy$$

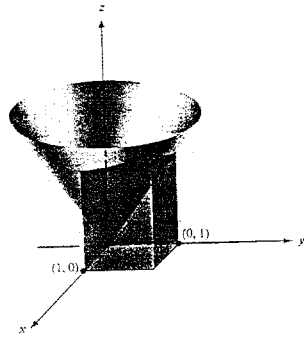
donde $R = [0, 1] \times [0, 1]$.

Solución:

Esta integral doble es igual al volumen de la región tridimensional mostrada en la figura. Como está planteada, es difícil evaluar esta integral. Como el integrando es una función sencilla de $r^2 = x^2 + y^2$, intentamos de nuevo un cambio de variables a coordenadas polares

$$T(r, \theta) = (r \cos \theta, r \sin \theta)$$

Esto nos llevará a simplificar el integrando aunque complicaremos la región de integración. Sin embargo, la simplificación es suficiente para permitirnos evaluar la integral.



Consideremos antes de todo las ecuaciones que describen los cuatro lados del cuadrado $R = [0, 1] \times [0, 1]$, para tener una indicación de donde van a variar las coordenadas polares r y θ .

$$x = 0 \rightarrow r \cos \theta = 0 \Rightarrow \begin{cases} r = 0 \text{ y } \theta \text{ cualquiera} \\ r \text{ cualquiera y } \theta = (2k + 1)\frac{\pi}{2}, \quad k \in \mathbb{Z}, \end{cases}$$

$$y = 0 \rightarrow r \sin \theta = 0 \Rightarrow \begin{cases} r = 0 \text{ y } \theta \text{ cualquiera} \\ r \text{ cualquiera y } \theta = k\pi, \quad k \in \mathbb{Z}, \end{cases}$$

$$x = 1 \rightarrow r \cos \theta = 1 \Rightarrow r = \frac{1}{\cos \theta} = \sec \theta$$

$$y = 1 \rightarrow r \sin \theta = 1 \Rightarrow r = \frac{1}{\sin \theta} = \operatorname{cosec} \theta.$$

Visto que las dos condiciones $x = 1$ e $y = 1$ nos dan curvas distintas para la coordenada polar r , es conveniente dividir el cuadrado en dos triángulos:

$$T_1 = \{(x, y) / 0 \leq x \leq 1, 0 \leq y \leq x\}, \quad T_2 = \{(x, y) / 0 \leq y \leq 1, 0 \leq x \leq y\}.$$

Luego, el primer triángulo nos dará

$$0 \leq x \leq 1 \Rightarrow 0 \leq r \cos \theta \leq 1 \Rightarrow 0 \leq r \leq \frac{1}{\cos \theta} = \sec \theta$$

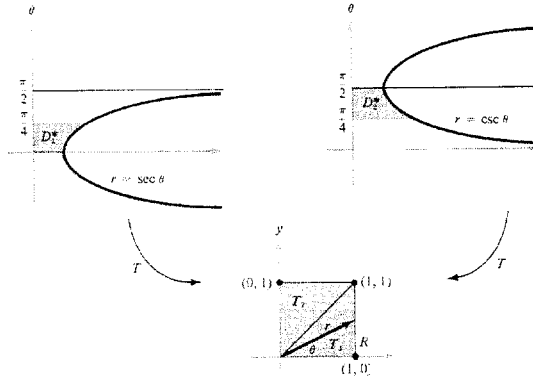
$$0 \leq y \leq x \Rightarrow 0 \leq r \sin \theta \leq r \cos \theta \Rightarrow 0 \leq \operatorname{tg} \theta \leq 1 \Rightarrow 0 \leq \theta \leq \frac{\pi}{4}.$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Para el segundo triángulo obtenemos

$$0 \leq y \leq 1 \Rightarrow 0 \leq r \operatorname{sen} \theta \leq 1 \Rightarrow 0 \leq r \leq \frac{1}{\operatorname{sen} \theta} = \operatorname{cosec} \theta$$

$$0 \leq x \leq y \Rightarrow 0 \leq r \cos \theta \leq r \operatorname{sen} \theta \Rightarrow 0 \leq \cotg \theta \leq 1 \Rightarrow \frac{\pi}{4} \leq \theta \leq \frac{\pi}{2}.$$



La transformación T es uno a uno excepto cuando $r = 0$, de modo que se puede aplicar el teorema 12.12. Por la simetría de $z = \sqrt{x^2 + y^2}$ en R , podemos escribir

$$\int_R \sqrt{x^2 + y^2} \, dx \, dy = 2 \int_{T_1} \sqrt{x^2 + y^2} \, dx \, dy.$$

Entonces, usando las dos siguientes integrales escogidas en tablas

$$\int \sec^3 \theta \, d\theta = \frac{\sec \theta \operatorname{tg} \theta}{2} + \frac{1}{2} \int \sec \theta \, d\theta,$$

$$\int \sec \theta \, d\theta = \frac{1}{2} \log |\sec \theta + \operatorname{tg} \theta|,$$

obtenemos

$$\begin{aligned} \int_R \sqrt{x^2 + y^2} \, dx \, dy &= 2 \int_{T_1} \sqrt{x^2 + y^2} \, dx \, dy = 2 \int_{D_1} \sqrt{r^2} \, r \, dr \, d\theta = 2 \int_{D_1} r^2 \, dr \, d\theta = \\ &= 2 \int_{D_1} r^2 \, dr \, d\theta = 2 \int_0^{\pi/4} \left[\int_0^{\operatorname{cosec} \theta} r^2 \, dr \right] d\theta = \frac{2}{3} \int_0^{\pi/4} \sec^3 \theta \, d\theta = \\ &= \frac{2}{3} \left[\left[\frac{\sec \theta \operatorname{tg} \theta}{2} \right]_0^{\pi/4} + \frac{1}{2} \int_0^{\pi/4} \sec \theta \, d\theta \right] = \frac{2}{3} \left[\frac{\sqrt{2}}{2} + \frac{1}{2} [\log |\sec \theta + \operatorname{tg} \theta|]_0^{\pi/4} \right] = \\ &= \frac{2}{3} \left[\frac{\sqrt{2}}{2} + \frac{1}{2} \log(1 + \sqrt{2}) \right] = \frac{1}{3} [\sqrt{2} + \log(1 + \sqrt{2})]. \end{aligned}$$

12.6.3 Cambio de variables en la integral triple.

Así como en el caso de $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ tenemos el teorema 12.12 que nos permite usar el cambio de variables para evaluar las integrales dobles, también hay una fórmula de cambio de variables para integrales triples. Antes de enunciar este nuevo teorema, debemos definir el jacobiano de una transformación de \mathbb{R}^3 a \mathbb{R}^3 , que es una extensión sencilla del caso de dos variables.

Definición. Sea $T: W \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ una función C^1 definida por las tres ecuaciones $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$. Entonces, el **jacobiano** de T es el determinante de la matriz derivada $DT(x, y)$ de T :

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}.$$

El valor absoluto de este determinante es igual al volumen (ver sección 9.3) del paralelepípedo determinado por los vectores

$$\begin{aligned} \vec{T}_u &= \frac{\partial x}{\partial u} \vec{i} + \frac{\partial y}{\partial u} \vec{j} + \frac{\partial z}{\partial u} \vec{k}, \\ \vec{T}_v &= \frac{\partial x}{\partial v} \vec{i} + \frac{\partial y}{\partial v} \vec{j} + \frac{\partial z}{\partial v} \vec{k}, \\ \vec{T}_w &= \frac{\partial x}{\partial w} \vec{i} + \frac{\partial y}{\partial w} \vec{j} + \frac{\partial z}{\partial w} \vec{k}. \end{aligned}$$

Así, como en el caso de dos variables, el jacobiano mide cómo la transformación T distorsiona su dominio. Para integrales triples, la fórmula del cambio de variables toma la forma dada en el siguiente teorema.

Teorema 12.13 Cambio de variables para integrales triples.

Sean $D = T(D^*)$ y D^* regiones elementales en los espacios xyz y uvw , respectivamente. Sea $T: D^* \rightarrow D$, la transformación que da el cambio de coordenadas

$$T(u, v, w) \rightsquigarrow (x = x(u, v, w), y = y(u, v, w), z = z(u, v, w)),$$

de clase C^1 y uno a uno en D^* , excepto quizá en un conjunto que sea la unión de gráficas de dos variables. Entonces, para cualquier función integrable $f: D \rightarrow \mathbb{R}$, tenemos

$$\begin{aligned} \int_D f(x, y, z) dx dy dz &= \\ &= \int_{D^*} f(x(u, v, w), y(u, v, w), z(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du dv dw. \end{aligned} \tag{12.15}$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Antes de ver algún ejemplo, apliquemos la fórmula (12.15) para coordenadas cilíndricas y esféricas (ver la sección 9.4). Para hacerlo necesitamos calcular el jacobiano de la función que define el cambio de coordenadas.

Las **coordenadas cilíndricas** (r, θ, z) de un punto (x, y, z) , ver la fórmula (9.2), vienen definidas por:

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

con $r \geq 0$, $0 \leq \theta < 2\pi$ y $-\infty < z < \infty$; luego el jacobiano será

$$\frac{\partial(x, y, z)}{\partial(r, \theta, z)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r.$$

La fórmula (12.15) en el caso de coordenadas cilíndricas, se escribe entonces como

$$\int_D f(x, y, z) dx dy dz = \int_{D^*} f(r \cos \theta, r \sin \theta, z) r dr d\theta dz. \quad (12.16)$$

Las coordenadas cilíndricas son particularmente útiles cuando existe un eje de simetría. Se suele tomar dicho eje como eje z .

Las **coordenadas esféricas** (ρ, θ, ϕ) de un punto (x, y, z) , ver la fórmula (9.3), vienen definidas por:

$$x = \rho \cos \theta \sin \phi, \quad y = \rho \sin \theta \sin \phi, \quad z = \rho \cos \phi,$$

con $\rho \geq 0$, $0 \leq \theta < 2\pi$ y $0 \leq \phi \leq \pi$; luego el jacobiano será

$$\begin{aligned} \frac{\partial(x, y, z)}{\partial(\rho, \theta, \phi)} &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \cos \theta \sin \phi & -\rho \sin \theta \sin \phi & \rho \cos \theta \cos \phi \\ \sin \theta \sin \phi & \rho \cos \theta \sin \phi & \rho \sin \theta \cos \phi \\ \cos \phi & 0 & -\rho \sin \phi \end{vmatrix} = \\ &= -\rho^2 \sin \phi. \end{aligned}$$

La fórmula (12.15) en el caso de coordenadas esféricas, se escribe entonces como

$$\begin{aligned} \int_D f(x, y, z) dx dy dz &= \\ &= \int_{D^*} f(\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi) \rho^2 \sin \phi d\rho d\theta d\phi. \end{aligned} \quad (12.17)$$

Las coordenadas esféricas se usan habitualmente para los problemas en los cuales hay un centro de simetría, que se suele tomar como el origen.

Para probar la validez de la fórmula (12.17), se debe mostrar que la transformación $S(\rho, \theta, \phi) = (x, y, z)$ en el conjunto D^* es uno a uno excepto en un conjunto que sea la unión de un número finito de gráficas de funciones continuas.

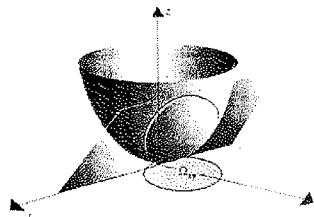
Ejemplo 1. Usa coordenadas cilíndricas para hallar el volumen del sólido T limitado superiormente por el plano $z = y$ e inferiormente por el paraboloides $z = x^2 + y^2$.

Solución:

En coordenadas cilíndricas

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

la ecuación del plano es $z = r \sin \theta$ y la del paraboloides es $z = r^2$. Resolviendo el sistema formado por estas dos ecuaciones, obtenemos $r = \sin \theta$. Esto significa que las dos superficies se cortan en una curva que está sobre el cilindro circular $r = \sin \theta$. La proyección de esta intersección sobre el plano xy es la circunferencia de ecuación $r = \sin \theta$.



La región de base consta de todos los puntos de \mathbb{R}^2 con coordenadas polares en el conjunto

$$0 \leq \theta \leq \pi, \quad 0 \leq r \leq \sin \theta.$$

Entonces, el sólido T es el conjunto de todos los puntos (x, y, z) con coordenadas cilíndricas en el conjunto T^*

$$0 \leq \theta \leq \pi, \quad 0 \leq r \leq \sin \theta, \quad r^2 \leq z \leq r \sin \theta.$$

Entonces, usando la fórmula (12.16) obtenemos

$$\begin{aligned} \int_T dx dy dz &= \int_{T^*} r dr d\theta dz = \int_0^\pi \int_0^{\sin \theta} \int_{r^2}^{r \sin \theta} r dz dr d\theta = \\ &= \int_0^\pi \int_0^{\sin \theta} (r^2 \sin \theta - r^3) dr d\theta = \int_0^\pi \left[\frac{r^3}{3} \sin \theta - \frac{r^4}{4} \right]_0^{\sin \theta} d\theta = \\ &= \frac{1}{12} \int_0^\pi \sin^4 \theta d\theta = \frac{1}{12} \cdot \frac{3}{8} \pi = \frac{1}{32} \pi. \end{aligned}$$

Para la última integral hemos usado las fórmulas

$$\int \sin^n \theta d\theta = -\frac{\sin^{n-1} \theta \cos \theta}{n} + \frac{n-1}{n} \int \sin^{n-2} \theta d\theta$$

y

$$\int \sin^2 m \theta d\theta = \frac{1}{2m} (m\theta - \sin m\theta \cos m\theta).$$

12.6. CAMBIO DE VARIABLES EN LA INTEGRAL MÚLTIPLE.

Ejemplo 2. Sea D la esfera de radio R y centro $(0, 0, 0)$ en \mathbb{R}^3 . Halla el volumen de D .

Solución:

El volumen de D es

$$\int_D dx dy dz.$$

Como vimos en la sección 12.5, esta integral se puede evaluar reduciéndola a integrales iteradas. También se puede calcular considerando D como un volumen de revolución. Pero aquí la evaluaremos usando coordenadas esféricas. Si D^* es la región tal que

$$0 \leq \rho \leq R, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \phi \leq \pi,$$

podemos usar la fórmula (12.17) y obtener

$$\begin{aligned} \int_D dx dy dz &= \int_0^\pi \int_0^{2\pi} \int_0^R \rho^2 \sin \phi d\rho d\theta d\phi = \frac{R^3}{3} \int_0^\pi \int_0^{2\pi} \sin \phi d\theta d\phi = \\ &= \frac{2\pi R^3}{3} \int_0^\pi \sin \phi d\phi = \frac{2\pi R^3}{3} [-\cos \phi]_0^\pi = \frac{4}{3} \pi R^3, \end{aligned}$$

que es la fórmula conocida para el volumen de una esfera sólida.

Ejemplo 3. Evalúa

$$\int_D \exp(x^2 + y^2 + z^2)^{3/2} dV,$$

donde D es la esfera unitaria en \mathbb{R}^3 .

Solución:

Notemos, antes de todo, que no es fácil integrar esta función usando integrales iteradas. Tratemos entonces el problema con un cambio de variables. Dado que el argumento de la función aparece como $x^2 + y^2 + z^2$, usando coordenadas esféricas, se puede reemplazar por una variable, ρ^2 . Como en el ejemplo anterior, consideremos D^* la región dada por

$$0 \leq \rho \leq 1, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \phi \leq \pi.$$

Podemos usar la fórmula (12.17) para obtener

$$\begin{aligned} \int_D \exp(x^2 + y^2 + z^2)^{3/2} dx dy dz &= \int_0^1 \int_0^\pi \int_0^{2\pi} e^{\rho^3} \rho^2 \sin \phi d\theta d\phi d\rho = \\ &= 2\pi \int_0^1 e^{\rho^3} \rho^2 \left[\int_0^\pi \sin \phi d\phi \right] d\rho = 2\pi \int_0^1 e^{\rho^3} \rho^2 [-\cos \phi]_0^\pi d\rho = \\ &= 4\pi \int_0^1 e^{\rho^3} \rho^2 d\rho = \frac{4}{3} \pi \int_0^1 e^{\rho^3} (3\rho^2) d\rho = \frac{4}{3} \pi [e^{\rho^3}]_0^1 = \frac{4}{3} \pi (e - 1). \end{aligned}$$

Ejercicios.

- 1) Transforma cada una de las integrales dadas en una o más integrales reiteradas en polares.

$$\int_0^1 \int_0^1 f(x, y) dy dx; \quad \int_0^1 \int_{1-x}^{\sqrt{1-x^2}} f(x, y) dy dx; \quad \int_0^1 \int_0^{x^2} f(x, y) dy dx.$$

- 2) Calcula cada una de las integrales dadas transformandola en coordenadas polares.

$$\begin{array}{ll} \text{a)} \int_0^{2a} \int_0^{\sqrt{2ax-x^2}} (x^2 + y^2) dy dx; & \text{b)} \int_0^1 \int_{x^2}^x \frac{1}{\sqrt{x^2 + y^2}} dy dx; \\ \text{c)} \int_0^a \int_0^x \sqrt{x^2 + y^2} dy dx; & \text{d)} \int_0^a \int_0^{\sqrt{a^2 - y^2}} (x^2 + y^2) dx dy; \\ \text{e)} \int_{-1}^1 \int_0^{\sqrt{1-y^2}} \sqrt{x^2 + y^2} dx dy; & \text{f)} \int_0^2 \int_0^{\sqrt{4-x^2}} \sqrt{x^2 + y^2} dy dx; \\ \text{g)} \int_{1/2}^1 \int_0^{\sqrt{1-x^2}} (x^2 + y^2)^{3/2} dy dx; & \text{h)} \int_0^{1/2} \int_0^{\sqrt{1-x^2}} xy \sqrt{x^2 + y^2} dy dx. \end{array}$$

- 3) Efectúa el cambio de variables $u = x+y$, $v = x-y$ en la integral $\int_0^1 \int_0^1 f(x, y) dx dy$.

- 4) Sea D la región $0 \leq y \leq x$ y $0 \leq x \leq 1$. Evalúa la integral $\int_D (x+y) dx dy$ haciendo el cambio de variables $x = u+v$, $y = u-v$. Verifica la respuesta obtenida evaluando directamente la integral, usando una integral iterada.

- 5) Sea $T(u, v) = T(x(u, v), y(u, v))$ la función definida por $T(u, v) = (4u, 2u + 3v)$. Sea D^* el rectángulo $[0, 1] \times [1, 2]$. Halla $D = T(D^*)$ y evalúa

$$\text{a)} \int_D xy dx dy; \quad \text{b)} \int_D (x-y) dx dy;$$

haciendo un cambio de variables para evaluarlas como integrales sobre D^* .

- 6) Sea $T(u, v) = T(x(u, v), y(u, v))$ la función definida por $T(u, v) = (u, v(1+u))$. Sea D^* el rectángulo $[0, 1] \times [1, 2]$. Halla $D = T(D^*)$ y evalúa

$$\text{a)} \int_D xy dx dy; \quad \text{b)} \int_D (x-y) dx dy;$$

haciendo un cambio de variables para evaluarlas como integrales sobre D^* .

- 7) Sea $T(u, v) = T(x(u, v), y(u, v))$ la función definida por $T(u, v) = (u^2 - v^2, 2uv)$. Sea D^* el conjunto de (u, v) con $u^2 + v^2 \leq 1$, $u \geq 0$, $v \geq 0$. Halla $D = T(D^*)$ y evalúa

Así mismo, para funciones de dos variables, la razón de la integral al área de D es

$$[f]_{\text{prom}} = \frac{\iint_D f(x, y) \, dx \, dy}{\iint_D dx \, dy}, \quad (12.18)$$

y se llama **valor promedio de f sobre D** .

Además, si D es la región del plano ocupado por una lámina plana que tiene una densidad de masa $f(x, y)$ en el punto $(x, y) \in D$ entonces definimos el **área** de la lámina como

$$A(D) = \iint_D dx \, dy; \quad (12.19)$$

la **masa** de la lámina como

$$m(D) = \iint_D f(x, y) \, dx \, dy; \quad (12.20)$$

y las **coordenadas del centro de gravedad** de la lámina. (\bar{x}, \bar{y}) , como

$$\bar{x} = \frac{\iint_D x f(x, y) \, dx \, dy}{m(D)} \quad \bar{y} = \frac{\iint_D y f(x, y) \, dx \, dy}{m(D)}. \quad (12.21)$$

Si L es una recta y $\delta(x, y)$ es la distancia del punto (x, y) de la lámina a la recta L , el **momento de inercia** de la lámina respecto a L es

$$I_L = \iint_D \delta(x, y)^2 f(x, y) \, dx \, dy.$$

En particular los momentos de inercia respecto a los ejes coordenados son

$$I_x = \iint_D y^2 f(x, y) \, dx \, dy \quad I_y = \iint_D x^2 f(x, y) \, dx \, dy. \quad (12.22)$$

Veamos algunos ejemplos.

Ejemplo 1. Halla el valor promedio de $f(x, y) = x \operatorname{sen}^2(xy)$ en $D = [0, \pi] \times [0, \pi]$.

Solución:

Primero calculamos

$$\begin{aligned} \iint_D f(x, y) \, dx \, dy &= \int_0^\pi \int_0^\pi x \operatorname{sen}^2(xy) \, dx \, dy = \int_0^\pi \left[\int_0^\pi x \frac{1 - \cos(2xy)}{2} \, dy \right] dx = \\ &= \int_0^\pi x \left[\frac{y}{2} - \frac{\operatorname{sen}(2xy)}{4x} \right]_0^\pi dx = \int_0^\pi \left(\frac{\pi x}{2} - \frac{\operatorname{sen}(2xy)}{4} \right) dx = \\ &= \left[\frac{\pi x^2}{4} - \frac{\cos(2xy)}{8\pi} \right]_0^\pi = \frac{\pi^3}{4} + \frac{\cos(2\pi^2) - 1}{8\pi}. \end{aligned}$$

12.7. ALGUNAS APLICACIONES DE LAS INTEGRALES MÚLTIPLES.

Usando la fórmula (12.18), tenemos

$$[f]_{\text{prom}} = \frac{\iint_D f(x, y) dx dy}{\iint_D dx dy} = \frac{\frac{\pi^3}{4} + \frac{\cos(2\pi^2) - 1}{8\pi}}{\pi^2} = \frac{\pi}{4} + \frac{\cos(2\pi^2) - 1}{8\pi^3} \approx 0.7839.$$

Ejemplo 2. Una lámina tiene forma de semidisco de radio a . Halla la masa de la lámina y su centro de masas dado que la densidad de la lámina varía proporcionalmente a la distancia del centro del lado recto de la lámina.

Solución:

Coloquemos la lámina sobre la región Ω : $-a \leq x \leq a$, $0 \leq y \leq \sqrt{a^2 - x^2}$. La densidad puede escribirse como

$$f(x, y) = k \sqrt{x^2 + y^2}.$$

Entonces, por la fórmula (12.20), y usando el cambio a coordenadas polares, tenemos

$$m(\Omega) = \iint_{\Omega} f(x, y) dx dy = \iint_{\Omega} k \sqrt{x^2 + y^2} dx dy = \int_0^{\pi} \int_0^a (kr) r dr d\theta = \frac{1}{3} k a^3 \pi.$$

Usamos ahora las fórmulas (12.21); para la primera coordenada obtenemos

$$\bar{x} = \frac{\iint_D x f(x, y) dx dy}{m(D)} = \frac{\iint_D x k \sqrt{x^2 + y^2} dx dy}{\frac{1}{3} k a^3 \pi} = \frac{0}{\frac{1}{3} k a^3 \pi} = 0,$$

por ser el integrando impar respecto de x ; y para la segunda, usando el cambio a coordenadas polares,

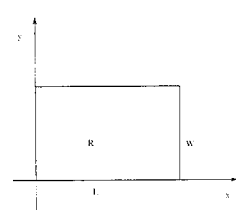
$$\begin{aligned} \bar{y} &= \frac{\iint_D y f(x, y) dx dy}{m(D)} = \frac{\iint_D y k \sqrt{x^2 + y^2} dx dy}{\frac{1}{3} k a^3 \pi} = \frac{\int_0^{\pi} \int_0^a r \operatorname{sen} \theta (kr) r dr d\theta}{\frac{1}{3} k a^3 \pi} = \\ &= \frac{k \left(\int_0^{\pi} \operatorname{sen} \theta d\theta \right) \left(\int_0^a r^3 dr \right)}{\frac{1}{3} k a^3 \pi} = \frac{\frac{1}{2} k a^4}{\frac{1}{3} k a^3 \pi} = \frac{3a}{2\pi}. \end{aligned}$$

Ejemplo 3. Una lámina rectangular de masa M , base L y altura W gira alrededor del lado menor izquierdo. Halla el momento de inercia de la lámina con respecto al eje si

- la lámina tiene densidad uniforme,
- la densidad varía proporcionalmente al cuadrado de la distancia al lado derecho.

Solución:

Situamos la lámina sobre el eje coordenado como en la figura, y llamamos R la región de base.



(a) Aquí, la densidad de masa es

$$f(x, y) = \frac{M}{LW},$$

luego usando la segunda de las fórmulas (12.22)

$$\begin{aligned} I_y &= \iint_R x^2 f(x, y) dx dy = \iint_R x^2 \frac{M}{LW} dx dy = \frac{M}{LW} \int_0^W \int_0^L x^2 dx dy = \\ &= \frac{M}{LW} W \int_0^L x^2 dx = \frac{M}{L} \frac{L^3}{3} = \frac{ML^2}{3}. \end{aligned}$$

(b) En este caso $f(x, y) = k(L - x)^2$; usando otra vez la segunda de las fórmulas (12.22) obtenemos

$$\begin{aligned} I_y &= \iint_R x^2 f(x, y) dx dy = \iint_R x^2 k(L - x)^2 dx dy = k \int_0^W \int_0^L (L - x)^2 x^2 dx dy = \\ &= kW \int_0^L (L^2 - 2Lx + x^2) x^2 dx = \frac{1}{30} k L^5 W. \end{aligned}$$

De manera muy parecida, para funciones de tres variables, la razón de la integral al volumen de W es

$$[f]_{\text{prom}} = \frac{\iiint_W f(x, y, z) dx dy dz}{\iiint_W dx dy dz}, \quad (12.23)$$

y se llama **valor promedio de f sobre W** .

Análogamente, si W es la región del espacio \mathbb{R}^3 ocupado por un sólido que tiene una densidad de masa $f(x, y, z)$ en el punto $(x, y, z) \in W$ entonces, el **volumen** del sólido es

$$\iiint_W dx dy dz; \quad (12.24)$$

y la **masa** del sólido es

$$m(W) = \iiint_W f(x, y, z) dx dy dz; \quad (12.25)$$

12.7. ALGUNAS APLICACIONES DE LAS INTEGRALES MÚLTIPLES.

y las **coordenadas del centro de gravedad** del sólido, $(\bar{x}, \bar{y}, \bar{z})$, vienen dadas por

$$\begin{aligned}\bar{x} &= \frac{\iiint_W x f(x, y, z) dx dy dz}{m(W)} \\ \bar{y} &= \frac{\iiint_W y f(x, y, z) dx dy dz}{m(W)} \\ \bar{z} &= \frac{\iiint_W z f(x, y, z) dx dy dz}{m(W)}.\end{aligned}\tag{12.26}$$

Si el objeto es homogéneo, es decir su densidad es constante e igual a $\frac{M}{V}$, el centro de gravedad sólo depende de la geometría del sólido y coincide con el centroide $(\bar{x}, \bar{y}, \bar{z})$ del espacio ocupado por el sólido. La densidad resulta entonces irrelevante y las **coordenadas del centroide** son

$$\begin{aligned}\bar{x} &= \frac{\iiint_W x dx dy dz}{V} \\ \bar{y} &= \frac{\iiint_W y dx dy dz}{V} \\ \bar{z} &= \frac{\iiint_W z dx dy dz}{V}.\end{aligned}\tag{12.27}$$

Si L es una recta en \mathbb{R}^3 y $\delta(x, y, z)$ es la distancia del punto (x, y, z) del sólido a la recta L , el **momento de inercia** del sólido respecto a L es

$$I_L = \iiint_W \delta(x, y, z)^2 f(x, y, z) dx dy dz.$$

En particular los momentos de inercia respecto a los ejes coordenados son

$$\begin{aligned}I_x &= \iiint_W (y^2 + z^2) f(x, y, z) dx dy dz \\ I_y &= \iiint_W (x^2 + z^2) f(x, y, z) dx dy dz \\ I_z &= \iiint_W (x^2 + y^2) f(x, y, z) dx dy dz.\end{aligned}\tag{12.28}$$

Si P es un plano en \mathbb{R}^3 y $\pi(x, y, z)$ es la distancia del punto (x, y, z) del sólido al plano P , el **momento de inercia** del sólido respecto al plano P es

$$I_P = \iiint_W \pi(x, y, z)^2 f(x, y, z) dx dy dz.$$

En particular los momentos de inercia respecto a los planos coordenados son

$$\begin{aligned} I_{xy} &= \iiint_W z^2 f(x, y, z) \, dx \, dy \, dz \\ I_{yz} &= \iiint_W x^2 f(x, y, z) \, dx \, dy \, dz \\ I_{xz} &= \iiint_W y^2 f(x, y, z) \, dx \, dy \, dz. \end{aligned} \quad (12.29)$$

Ejemplo 1. Halla la masa de un cilindro circular recto de radio R y altura h , siendo la densidad proporcional a la distancia desde una de las bases.

Solución:

Llamemos T a dicho sólido. Podemos caracterizar T por las siguientes desigualdades

$$-R \leq x \leq R, \quad -\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}, \quad 0 \leq z \leq h.$$

Suponiendo que la densidad es proporcional a la distancia a la base inferior, tenemos $f(x, y, z) = kz$. Luego, usando propiedades de simetría del cilindro o directamente usando la fórmula conocida del área de un círculo, obtenemos

$$\begin{aligned} M &= \iiint_T kz \, dx \, dy \, dz = k \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \int_0^h z \, dz \, dy \, dx = k \frac{h^2}{2} \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \, dx = \\ &= k \frac{h^2}{2} \int_{-R}^R \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \, dx = k \frac{h^2}{2} (\text{área de base}) = 4k \frac{h^2}{2} \int_0^R \int_0^{\sqrt{R^2-x^2}} dy \, dx = \\ &= 2kh^2 \int_0^R \sqrt{R^2 - x^2} \, dx = 2kh^2 \frac{\pi R^2}{4} = \frac{1}{2} \pi kh^2 R^2. \end{aligned}$$

Para la última integral hemos usado la fórmula

$$\int \sqrt{a^2 - x^2} \, dx = \frac{x}{2} \sqrt{a^2 - x^2} + \frac{a^2}{2} \arcsen \frac{x}{a}.$$

Ejemplo 2. Halla la masa de un cilindro circular recto de radio R y altura h , siendo la densidad proporcional a la distancia al eje del cilindro.

Solución:

Situemos el cilindro T sobre el plano xy de tal modo que el eje z coincida con el eje de T . Como en el ejemplo anterior podemos caracterizar T usando coordenadas cartesianas, sin embargo en este caso tenemos que la densidad es proporcional a la distancia al eje del cilindro, luego, será $f(x, y, z) = k\sqrt{x^2 + y^2}$. Resulta entonces conveniente usar coordenadas cilíndricas y describir T como los puntos (x, y, z) cuyas coordenadas cilíndricas están en el conjunto T^* descrito por las siguientes desigualdades

$$0 \leq r \leq R, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq z \leq h.$$

12.7. ALGUNAS APLICACIONES DE LAS INTEGRALES MÚLTIPLES.

Luego,

$$\begin{aligned} M &= \iiint_T k \sqrt{x^2 + y^2} dx dy dz = \iiint_{T^*} (k r) r dr d\theta dz = k \int_0^R \int_0^{2\pi} \int_0^h r^2 dz d\theta dr = \\ &= 2\pi h k \int_0^R r^2 dr = \frac{2}{3} \pi h k R^3. \end{aligned}$$

Ejemplo 3. Localiza el centroide del sólido T limitado superiormente por el plano $z = y$ e inferiormente por el paraboloides $z = x^2 + y^2$.

Solución:

En un ejemplo de la sección anterior, usamos coordenadas cilíndricas para hallar el volumen de dicho sólido T , obteniendo

$$V = \int_T dx dy dz = \int_{T^*} r dr d\theta dz = \frac{1}{32} \pi.$$

siendo T^* el conjunto de todos los puntos con coordenadas cilíndricas

$$0 \leq \theta \leq \pi, \quad 0 \leq r \leq \sin \theta, \quad r^2 \leq z \leq r \sin \theta.$$

Usamos ahora las ecuaciones (12.27) para hallar las coordenadas $(\bar{x}, \bar{y}, \bar{z})$.

Primero notemos que T es simétrico con respecto al plano yz , y entonces resulta que $\bar{x} = 0$. Para hallar \bar{y} calculemos

$$\iiint_T y dx dy dz = \iiint_{T^*} (r \sin \theta) r dr d\theta dz = \int_0^\pi \int_0^{\sin \theta} \int_{r^2}^{r \sin \theta} r^2 \sin \theta dr d\theta dz = \frac{\pi}{64};$$

$$\text{luego, por la segunda de (12.27): } \bar{y} = \frac{\iiint_T y dx dy dz}{V} = \frac{\frac{\pi}{64}}{\frac{\pi}{32}} = \frac{1}{2}.$$

Para la tercera coordenada \bar{z} , obtenemos

$$\iiint_T z dx dy dz = \iiint_{T^*} z r dr d\theta dz = \int_0^\pi \int_0^{\sin \theta} \int_{r^2}^{r \sin \theta} z r \sin \theta dr d\theta dz = \frac{5\pi}{384};$$

$$\text{luego, por la tercera de (12.27): } \bar{z} = \frac{\iiint_T z dx dy dz}{V} = \frac{\frac{5\pi}{384}}{\frac{\pi}{32}} = \frac{5}{12}.$$

Luego el centroide es $(\bar{x}, \bar{y}, \bar{z}) = \left(0, \frac{1}{2}, \frac{5}{12}\right)$.

Ejemplo 4. Calcula el momento de inercia I_z del sólido por encima del plano xy acotado por el paraboloides $z = x^2 + y^2$ y el cilindro $x^2 + y^2 = a^2$, suponiendo que a y la densidad de masa son constantes.

Solución:

El paraboloide y el cilindro se intersectan en el plano $z = a^2$. Para describir este sólido es conveniente usar coordenadas cilíndricas y describir T como los puntos (x, y, z) cuyas coordenadas cilíndricas están en el conjunto T^* descrito por las siguientes desigualdades

$$0 \leq r \leq a, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq z \leq r^2.$$

Dado que se ha supuesto la densidad constante, pondremos $f(x, y, z) = \rho = \text{constante}$. Finalmente, usando la tercera de las fórmulas (12.28), obtenemos

$$I_z = \iiint_{T^*} r^2 \rho r \, dr \, d\theta \, dz = \rho \int_0^a \int_0^{2\pi} \int_0^{r^2} r^3 \, dz \, d\theta \, dr = \frac{1}{3} \pi \rho a^6.$$

Ejercicios.

- 1) Halla el centro de masa del rectángulo $[0, 1] \times [0, 1]$ si la densidad de masa es e^{x+y} .
- 2) El cubo $[1, 2] \times [1, 2] \times [1, 2]$ tiene densidad de masa $\rho(x, y, z) = (1+x)e^z y$. Halla la masa de dicho cubo.
- 3) Halla el centro de masa de la región semiesférica W definida por las desigualdades $x^2 + y^2 + z^2 \leq 1, z \geq 0$, suponiendo la densidad constante.
- 4) La temperatura en los puntos del cubo $[-1, 1] \times [-1, 1] \times [-1, 1]$ es proporcional al cuadrado de la distancia al origen.
 - a) ¿Cuál es la temperatura promedio?
 - b) ¿En qué puntos del cubo la temperatura es igual a la temperatura promedio?
- 5) Calcula la masa M de una esfera sólida de radio 1 sabiendo que la densidad es proporcional al cuadrado de la distancia a su centro.

Capítulo 13

Integrales sobre trayectorias y superficies.

En el capítulo anterior hemos estudiado la integración sobre regiones en \mathbb{R}^2 y \mathbb{R}^3 . En este capítulo estudiaremos la integración a lo largo de curvas y la integración sobre superficies. En el núcleo de este tema se encuentran tres grandes teoremas relativos a la integración: el teorema de Green, el teorema de Stokes y el teorema de Gauss (usualmente conocido como teorema de la divergencia).

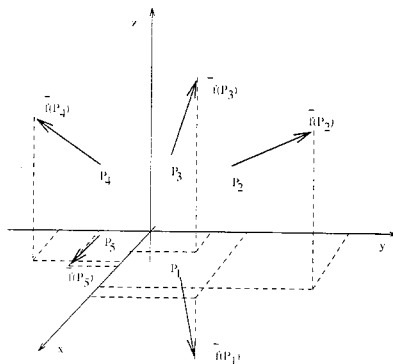
En última instancia, estos tres teoremas están basados en el teorema fundamental del cálculo integral y todos pueden enunciarse de una forma similar:

“una integral sobre un conjunto S = a una integral relacionada sobre la frontera de S .”

13.1 Campos vectoriales.

Definición. Un **campo vectorial** en \mathbb{R}^n es una función $\vec{F}: A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ que asigna a cada punto \vec{x} en su dominio A un vector $\vec{F}(\vec{x})$.

Podemos ilustrar gráficamente \vec{F} asignando una flecha a cada punto.



Es conveniente trazar la flecha que representa $\vec{F}(\vec{x})$ de modo que comience en \vec{x} , no en el origen.

De manera análoga, una función $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ que asigna un número a cada punto se llama **campo escalar**. Por ejemplo, un campo vectorial $\vec{F}(x, y, z)$ en \mathbb{R}^3 tiene tres campos escalares componentes F_1, F_2, F_3 , de modo que

$$\vec{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)).$$

Si cada campo componente es una función de clase C^k , decimos que el campo vectorial \vec{F} es de clase C^k . Supondremos que los campos vectoriales con los cuales trabajaremos son de clase C^1 .

Ejemplo 1. Consideremos una pieza de material que se calienta por un lado y se enfría por otro. La temperatura en cada punto dentro del cuerpo produce un campo escalar $T(x, y, z)$. El flujo real de calor se puede marcar mediante un campo de flechas que indiquen la dirección y magnitud del flujo. Esta energía o **campo vectorial de flujo de calor** está dado por $\vec{J} = -k \vec{\nabla} T$, donde $k > 0$ es una constante llamada conductividad y $\vec{\nabla} T$ es el gradiente de la función con valores reales T . El calor fluye de las regiones calientes hacia las frías, pues $-\vec{\nabla} T$ apunta en la dirección hacia donde T decrece (ver la sección 10.5).

Ejemplo 2. La fuerza de atracción de la Tierra sobre una masa m puede describirse mediante un campo vectorial en \mathbb{R}^3 , el campo de fuerza gravitacional. De acuerdo con la ley de Newton, este campo está dado por

$$\vec{F} = -G \frac{mM}{r^3} \vec{r},$$

donde $\vec{r}(x, y, z) = (x, y, z)$ y $r = \|\vec{r}\|$. \vec{F} es en realidad un **campo gradiente** $\vec{F} = -\vec{\nabla} V$, con $V = -G \frac{mM}{r}$. \vec{F} apunta en la dirección en que V decrece, es decir hacia el centro de la Tierra.

En general, un campo vectorial no tiene que ser un campo gradiente. Sin embargo, el concepto de superficie equipotencial tiene sentido sólo si el campo vectorial resulta ser un campo gradiente.

Nos dedicaremos en el resto de esta sección a estudiar los campos gradientes, y a presentar técnicas que volveremos a usar en secciones sucesivas.

Empezamos mostrando, con algunos ejemplos, cómo hallar la función $f(x, y)$ dado su gradiente

$$\vec{\nabla} f(x, y) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j}.$$

Ejemplo 1. Halla f sabiendo que

$$\vec{\nabla} f(x, y) = y^2 \vec{i} + (2xy - 1) \vec{j}.$$

13.1. CAMPOS VECTORIALES.

Solución:

Dado que

$$\vec{\nabla} f(x, y) = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j}.$$

tenemos

$$\frac{\partial f}{\partial x} = y^2, \quad \frac{\partial f}{\partial y} = 2xy - 1.$$

Integrando la primera derivada parcial $\frac{\partial f}{\partial x}$ con respecto de la variable x , tenemos

$$f(x, y) = xy^2 + \phi(y)$$

donde $\phi(y)$ es independiente de x pero puede depender de y . La diferenciación respecto de y da

$$\frac{\partial f}{\partial y} = 2xy + \phi'(y).$$

Las dos ecuaciones para la derivada parcial con respecto de y , $\frac{\partial f}{\partial y}$, se pueden cumplir solamente cuando

$$\phi'(y) = -1 \quad \text{luego} \quad \phi(y) = -y + c \quad \Rightarrow \quad f(x, y) = x^2y - y + c.$$

Ejemplo 2. Halla f sabiendo que

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

Solución:

Aquí tenemos que

$$\frac{\partial f}{\partial x} = \sqrt{y} - \frac{y}{2\sqrt{x}} + 2x, \quad \frac{\partial f}{\partial y} = \frac{x}{2\sqrt{y}} - \sqrt{x} + 1.$$

Como en el ejemplo anterior, integrando la primera derivada parcial $\frac{\partial f}{\partial x}$ con respecto de la variable x , tenemos

$$f(x, y) = x\sqrt{y} - \sqrt{x}y + x^2 + \phi(y)$$

donde $\phi(y)$ es independiente de x pero puede depender de y . La diferenciación respecto de y da

$$\frac{\partial f}{\partial y} = \frac{x}{2\sqrt{y}} - \sqrt{x} + \phi'(y).$$

Las dos ecuaciones para la derivada parcial $\frac{\partial f}{\partial y}$ pueden cumplirse solamente cuando

$$\phi'(y) = 1 \quad \text{luego} \quad \phi(y) = y + c \quad \Rightarrow \quad f(x, y) = x\sqrt{y} - \sqrt{x}y + x^2 + y + c.$$

Nótese que en ambos ejemplos, podríamos haber empezado integrando $\frac{\partial f}{\partial y}$ respecto de la variable y y luego diferenciando respecto de la variable x . Este procedimiento nos habría llevado al mismo resultado.

Veamos ahora con un ejemplo, que no todas las combinaciones lineales del tipo

$$P(x, y)\vec{i} + Q(x, y)\vec{j}$$

son gradientes de alguna función escalar $f(x, y)$.

Ejemplo. Demuestra que el campo

$$y\vec{i} - x\vec{j}$$

no proviene de un gradiente.

Solución:

Supongamos que se puede hallar mediante un gradiente. Entonces existe una función $f(x, y)$ tal que

$$\vec{\nabla} f(x, y) = y\vec{i} - x\vec{j}.$$

Evidentemente sería

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= y, & \frac{\partial f}{\partial y}(x, y) &= -x, \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= 1, & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= -1. \end{aligned}$$

Pero esto contradice el teorema 10.10: las cuatro derivadas parciales en consideración son todas continuas, en consecuencia, de acuerdo con el teorema 10.10, hemos de tener

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

Esta contradicción demuestra que el campo vectorial $y\vec{i} - x\vec{j}$ no proviene de un gradiente.

Abordemos ahora el problema de reconocer qué combinaciones lineales del tipo

$$P(x, y)\vec{i} + Q(x, y)\vec{j} \tag{13.1}$$

son gradientes de alguna función escalar $f(x, y)$.

Es posible demostrar que si $P(x, y)$ y $Q(x, y)$ son dos funciones de dos variables, ambas continuamente diferenciables en un abierto simplemente conexo Ω , entonces la combinación lineal $P(x, y)\vec{i} + Q(x, y)\vec{j}$, es un gradiente si y sólo si se satisface la condición

$$\frac{\partial P}{\partial y}(x, y) = \frac{\partial Q}{\partial x}(x, y) \tag{13.2}$$

para todo (x, y) en el abierto Ω (para una demostración alternativa ver el teorema 13.7).

13.1. CAMPOS VECTORIALES.

(Se dice que un abierto es conexo si cualquier par de puntos del conjunto se pueden unir por un camino poligonal que esté enteramente en el conjunto; se dice simplemente conexo si dada una curva cerrada C contenida en el abierto, todo el interior de la curva está también en el abierto).

Es fácil ver que si suponemos que el campo es un gradiente

$$P(x, y) \vec{i} + Q(x, y) \vec{j} = \vec{\nabla} f(x, y) = \frac{\partial f}{\partial x}(x, y) \vec{i} + \frac{\partial f}{\partial y}(x, y) \vec{j}.$$

entonces podemos identificar

$$P(x, y) = \frac{\partial f}{\partial x}(x, y) \quad y \quad Q(x, y) = \frac{\partial f}{\partial y}(x, y).$$

Puesto que P y Q tienen derivadas parciales de primer orden continuas, f tiene derivadas parciales de segundo orden continuas. Luego, de acuerdo con el teorema 10.10 las derivadas parciales mixtas son iguales y tenemos

$$\frac{\partial P}{\partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial Q}{\partial x}(x, y).$$

La demostración de la segunda parte es algo más complicada, y no la presentamos.

Ejemplo 1. Discute si las funciones vectoriales siguientes son campos vectoriales gradientes

$$\vec{F}(x, y) = 2x \operatorname{sen} y \vec{i} + x^2 \cos y \vec{j} \quad y \quad \vec{F}(x, y) = xy \vec{i} + \frac{1}{2}(x+1)^2 y^2 \vec{j}.$$

Solución:

Ambas funciones vectoriales están definidas en todo punto.

La primera es el gradiente de una función escalar definida en todo punto, porque

$$\frac{\partial P}{\partial y}(x, y) = 2x \cos y = \frac{\partial Q}{\partial x}(x, y), \quad \text{para todo } (x, y).$$

La segunda no es un gradiente porque

$$\frac{\partial P}{\partial y}(x, y) = x \neq \frac{\partial Q}{\partial x}(x, y) = (x-1)y^2.$$

Ejemplo 2. La función vectorial

$$\vec{F}(x, y) = \frac{y}{x^2 + y^2} \vec{i} - \frac{x}{x^2 + y^2} \vec{j}$$

definida en el disco perforado $D = \{(x, y) / 0 < x^2 + y^2 < 1\}$ satisface

$$\frac{\partial P}{\partial y}(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2} = \frac{\partial Q}{\partial x}(x, y), \quad \forall (x, y) \in D.$$

Sin embargo, como veremos en un ejemplo de aplicación del teorema de Green en la sección 13.7, \vec{F} no es un gradiente en dicho conjunto, que no es simplemente conexo.

Finalmente, sólo nombraremos otro concepto importante, que es el de línea de flujo.

Definición. Si \vec{F} es un campo vectorial, una **línea de flujo** para \vec{F} es una trayectoria $\vec{\sigma}(t)$ tal que

$$\vec{\sigma}'(t) = \vec{F}(\vec{\sigma}(t)).$$

Esto es, \vec{F} produce el campo de velocidad de la trayectoria $\vec{\sigma}(t)$.

Geoméricamente, el problema de hallar una línea de flujo que pase por un punto dado \vec{x}_0 para un campo vectorial \vec{F} dado, es el de ensartar una curva por el campo vectorial de manera que el vector tangente a la curva coincida con el campo vectorial.

Analíticamente, el problema de hallar una línea de flujo que pase por un punto dado \vec{x}_0 en el tiempo $t = 0$ implica resolver la ecuación diferencial con condición inicial \vec{x}_0 ; esto es

$$\vec{\sigma}'(t) = \vec{F}(\vec{\sigma}(t)), \quad \vec{\sigma}(0) = \vec{x}_0.$$

Ejemplo. En el plano, el campo vectorial \vec{V} definido por

$$\vec{V}(x, y) = \frac{y}{x^2 + y^2} \vec{i} - \frac{x}{x^2 + y^2} \vec{j}$$

está definido en todo \mathbb{R}^2 menos el origen. Este es el campo de velocidad que aproxima al campo de velocidad del agua en movimiento "circular" tal como ocurre, por ejemplo, cuando se quita el tapón de una tina de agua. Las líneas de flujo en este caso son círculos.

Nótese que campos vectoriales diferentes pueden tener líneas de flujo que sean la misma curva geométrica.

13.1.1 Divergencia y rotacional de un campo vectorial.

Definición. Dado un campo vectorial C^1 , \vec{F} en \mathbb{R}^3

$$\vec{F} = (F_1, F_2, F_3) = F_1 \vec{i} + F_2 \vec{j} + F_3 \vec{k}.$$

la operación **divergencia** asocia a cada campo vectorial, la función de valores reales $\text{div } \vec{F} = \vec{\nabla} \cdot \vec{F}$ definida como sigue:

$$\text{div } \vec{F} = \vec{\nabla} \cdot \vec{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}. \quad (13.3)$$

donde el operador $\vec{\nabla}$ se define como $\vec{\nabla} = \frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$. En notación de operador la divergencia $\text{div } \vec{F}$ es el producto escalar de $\vec{\nabla}$ y \vec{F} .

13.1. CAMPOS VECTORIALES.

Habíamos ya encontrado el operador $\vec{\nabla}$ operando sobre una función escalar f para darnos el **gradiente** de f :

$$\vec{\nabla} f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k}. \quad (13.4)$$

Definición. La operación **rotacional** asocia a cada campo vectorial \vec{F} , el campo vectorial $\text{rot } \vec{F} = \vec{\nabla} \times \vec{F}$ definido como sigue:

$$\begin{aligned} \text{rot } \vec{F} &= \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \vec{i} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) \vec{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \vec{k}. \end{aligned} \quad (13.5)$$

En notación de operador el rotacional, $\text{rot } \vec{F}$, es el producto vectorial de $\vec{\nabla}$ y \vec{F} . Nótese que $\text{rot } \vec{F}$ es de clase C^{k-1} si \vec{F} es de clase C^k .

Supongamos conocidos la divergencia y el rotacional de un campo. **¿Qué nos dice esto acerca de dicho campo?** Para una respuesta definitiva, hemos de esperar al teorema de divergencia (teorema de Gauss) y al teorema de Stokes, pero, como anticipo, podemos adelantar alguna respuesta, aunque esta no sea la mejor posible. Consideremos \vec{v} como la velocidad de un fluido. La divergencia de \vec{v} en un punto P nos da una indicación de cómo el fluido tiende a acumularse en la proximidad de P , si hay divergencia negativa, o a alejarse de P , si hay divergencia positiva. Es decir, $\text{div } \vec{F}$ representa la tasa de expansión por unidad de volumen del fluido. Si $\text{div } \vec{F}$ es positivo, el fluido se está expandiendo; y si $\text{div } \vec{F}$ es negativo significa que el fluido se comprime. Si $\text{div } \vec{F} = 0$ en P significa físicamente que el fluido no se comprime o es **incompresible** en P . El rotacional en P mide la tendencia rotacional del fluido. Si $\text{rot } \vec{F} = \vec{0}$ en P significa físicamente que el fluido no tiene rotaciones o es **irrotacional** en P ; esto es, no tiene remolinos. Es decir, si colocamos en el fluido una pequeña rueda con aspas, esta se moverá con el fluido, pero no girará alrededor de su eje.

Ejemplo 1. Sea

$$\vec{v}(x, y, z) = \alpha x \vec{i} + \alpha y \vec{j} + \alpha z \vec{k}.$$

Halla la divergencia y el rotacional y da una interpretación de los mismos.

Solución:

Para la divergencia, usando la ecuación (13.3), obtenemos

$$\text{div } \vec{F} = \vec{\nabla} \cdot \vec{F} = \frac{\partial(\alpha x)}{\partial x} + \frac{\partial(\alpha y)}{\partial y} + \frac{\partial(\alpha z)}{\partial z} = 3\alpha.$$

y para el rotacional, usando la ecuación (13.5), obtenemos

$$\begin{aligned} \operatorname{rot} \vec{F} &= \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \alpha x & \alpha y & \alpha z \end{vmatrix} = \alpha \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = \\ &= \left(\frac{\partial z}{\partial y} - \frac{\partial y}{\partial z} \right) \vec{i} - \left(\frac{\partial z}{\partial x} - \frac{\partial x}{\partial z} \right) \vec{j} + \left(\frac{\partial y}{\partial x} - \frac{\partial x}{\partial y} \right) \vec{k} = \vec{0}. \end{aligned}$$

El campo de este ejemplo

$$\vec{v}(x, y, z) = \alpha x \vec{i} + \alpha y \vec{j} + \alpha z \vec{k} = \alpha \vec{r}$$

puede considerarse como un campo de velocidades de un fluido con movimiento radial: hacia el origen si $\alpha < 0$ y alejándose del origen si $\alpha > 0$.

En la figura mostramos un punto (x, y, z) , un entorno esférico de este punto, y un cono con vértice en el origen que es tangente a la frontera de este entorno.



Tenemos que hacer notar dos cosas:

1. todo el fluido que está en el cono permanece en el cono.
2. la velocidad del fluido es proporcional a la distancia al origen.

Luego, si la divergencia 3α es negativa, el movimiento se produce hacia el origen y la cantidad de fluido presente en el entorno crece dado que el fluido que entra se mueve más rápidamente que el fluido que sale, además el área de entrada es mayor que el área de salida. Sin embargo, si la divergencia 3α es positiva, el movimiento que se produce es de alejamiento del origen, luego el entorno pierde fluido ya que el fluido que entra se mueve más lentamente que el fluido que sale, y el área de entrada es menor que el área de salida. Dado que el movimiento es radial, el fluido carece de tendencia rotacional, y como cabe esperar, su rotacional es nulo.

Ejemplo 2. Sea, con ω una constante positiva.

$$\vec{v}(x, y, z) = -\omega y \vec{i} + \omega x \vec{j}.$$

Halla la divergencia y el rotacional y da una interpretación de los mismos.

Solución:

Para la divergencia, usando la ecuación (13.3), obtenemos

$$\operatorname{div} \vec{F} = \vec{\nabla} \cdot \vec{F} = \frac{\partial(-\omega y)}{\partial x} + \frac{\partial(\omega x)}{\partial y} + \frac{\partial(0)}{\partial z} = 0.$$

13.1. CAMPOS VECTORIALES.

y para el rotacional, usando la ecuación (13.5), obtenemos

$$\begin{aligned} \operatorname{rot} \vec{F} &= \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix} = \\ &= \left(\frac{\partial(0)}{\partial y} - \frac{\partial(\omega x)}{\partial z} \right) \vec{i} - \left(\frac{\partial(0)}{\partial x} - \frac{\partial(-\omega y)}{\partial z} \right) \vec{j} + \left(\frac{\partial(\omega x)}{\partial x} - \frac{\partial(-\omega y)}{\partial y} \right) \vec{k} = \\ &= 2\omega \vec{k}. \end{aligned}$$

El campo de este ejemplo es el campo de velocidades de una rotación uniforme (un fluido o una rueda), con velocidad angular ω , en sentido contrario al de las agujas del reloj alrededor del eje z . Esto se puede comprobar observando que \vec{v} es perpendicular a \vec{r} :

$$\vec{v} \cdot \vec{r} = (-\omega y \vec{i} + \omega x \vec{j} + 0 \vec{k}) \cdot (x \vec{i} + y \vec{j} + z \vec{k}) = -\omega y x + \omega x y = 0$$

y la velocidad en cada punto es ωR donde R es el radio de rotación:

$$v = \sqrt{\omega^2 y^2 + \omega^2 x^2} = \omega \sqrt{x^2 + y^2} = \omega R.$$

¿Qué relación existe entre el rotacional $2\omega \vec{k}$ y la rotación realizada? El vector velocidad angular es $\vec{\omega} = \omega \vec{k}$; tal vector se define como el vector de longitud ω , paralelo al eje de rotación orientado de tal forma que, visto desde el extremo de $\vec{\omega}$, el fluido o la rueda gire en el sentido contrario al de las agujas del reloj. Luego, en este caso, el rotacional es dos veces el vector velocidad angular.

Con esta rotación ni crece ni decrece la cantidad de fluido en ningún entorno, y como hemos visto, la divergencia es idénticamente nula.

Los teoremas siguientes enuncian dos relaciones básicas entre el gradiente y el rotacional, el primero, y entre la divergencia y el rotacional, el segundo.

Teorema 13.1 Para cualquier función f de clase \mathcal{C}^2 , tenemos

$$\vec{\nabla} \times (\vec{\nabla} f) = \vec{0},$$

esto es, el rotacional de cualquier gradiente es el vector nulo.

Demostración:

Por la fórmula (13.4) tenemos

$$\vec{\nabla} f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k}$$

y, por la definición (13.5) del rotacional tenemos

$$\begin{aligned} \operatorname{rot} \vec{F} &= \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} = \\ &= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y} \right) \vec{i} - \left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z} \right) \vec{j} + \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) \vec{k}. \end{aligned}$$

Cada componente es nula debido a las propiedades de simetría de las derivadas parciales mixtas (ver teorema 10.10).

c.q.d.

Teorema 13.2 Para cualquier campo vectorial \vec{F} de clase C^2 , tenemos

$$\operatorname{div} \operatorname{rot} \vec{F} = \vec{\nabla} \cdot (\vec{\nabla} \times \vec{F}) = 0.$$

esto es, la divergencia de cualquier rotacional es cero.

Demostración:

Por la fórmula (13.5) del rotacional tenemos

$$\begin{aligned} \operatorname{rot} \vec{F} &= \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \vec{i} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) \vec{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \vec{k}. \end{aligned}$$

Entonces, por la fórmula (13.3) de la divergencia y por las propiedades de simetría de las derivadas parciales mixtas, obtenemos

$$\operatorname{div} \operatorname{rot} \vec{F} = \frac{\partial}{\partial x} \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) - \frac{\partial}{\partial y} \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) + \frac{\partial}{\partial z} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) = 0.$$

c.q.d.

Definición. El **operador de Laplace** ∇^2 , que opera sobre funciones f , está definido como sigue

$$\nabla^2 f = \vec{\nabla} \cdot (\vec{\nabla} f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}. \quad (13.6)$$

Si $\vec{F} = (F_1, F_2, F_3)$ es un campo vectorial C^2 , también podemos definir $\nabla^2 \vec{F}$ en términos de componentes

$$\nabla^2 \vec{F} = (\nabla^2 F_1, \nabla^2 F_2, \nabla^2 F_3). \quad (13.7)$$

13.1. CAMPOS VECTORIALES.

Este operador juega un papel importante en muchas leyes físicas.

Ejemplo 1. Si $f(x, y, z) = x^2 + y^2 + z^2$, usando la fórmula (13.6), tenemos

$$\begin{aligned}\nabla^2 f &= \frac{\partial^2 (x^2 + y^2 + z^2)}{\partial x^2} + \frac{\partial^2 (x^2 + y^2 + z^2)}{\partial y^2} + \frac{\partial^2 (x^2 + y^2 + z^2)}{\partial z^2} = \\ &= 2 + 2 + 2 = 6.\end{aligned}$$

Ejemplo 2. Si $f(x, y, z) = e^{xyz}$, usando la fórmula (13.6), tenemos

$$\begin{aligned}\nabla^2 f &= \frac{\partial^2 (e^{xyz})}{\partial x^2} + \frac{\partial^2 (e^{xyz})}{\partial y^2} + \frac{\partial^2 (e^{xyz})}{\partial z^2} = \\ &= \frac{\partial (yz e^{xyz})}{\partial x} + \frac{\partial (xz e^{xyz})}{\partial y} + \frac{\partial (xz e^{xyz})}{\partial z} = \\ &= y^2 z^2 e^{xyz} + x^2 z^2 e^{xyz} + x^2 z^2 e^{xyz} = \\ &= (y^2 z^2 + x^2 z^2 + x^2 z^2) e^{xyz}.\end{aligned}$$

Ejercicios.

1) Halla la divergencia $\vec{\nabla} \cdot \vec{v}$ y el rotacional $\vec{\nabla} \times \vec{v}$ para los siguientes campos vectoriales.

a) $\vec{v}(x, y) = x \vec{i} + y \vec{j}$;

b) $\vec{v}(x, y) = y \vec{i} + x \vec{j}$;

c) $\vec{v}(x, y) = \frac{x}{x^2 + y^2} \vec{i} + \frac{y}{x^2 + y^2} \vec{j}$;

d) $\vec{v}(x, y) = \frac{y}{x^2 + y^2} \vec{i} + \frac{x}{x^2 + y^2} \vec{j}$;

e) $\vec{v}(x, y, z) = x \vec{i} + 2y \vec{j} + 3z \vec{k}$;

f) $\vec{v}(x, y, z) = yz \vec{i} + xz \vec{j} + xy \vec{k}$;

g) $\vec{v}(x, y, z) = x^2 y \vec{i} + y^2 z \vec{j} + x y^2 \vec{k}$;

h) $\vec{v}(x, y, z) = xyz \vec{i} + xz \vec{j} + z \vec{k}$;

i) $\vec{v}(\vec{r}) = e^x \vec{r}$;

l) $\vec{v}(\vec{r}) = r^2 \vec{r}$;

m) $\vec{v}(\vec{r}) = e^{r^2} (\vec{i} + \vec{j} + \vec{k})$;

n) $\vec{v}(\vec{r}) = e^{y^2} \vec{i} + e^{z^2} \vec{j} + e^{x^2} \vec{k}$.

2) Calcula el laplaciano $\nabla^2 f$ para los siguientes campos escalares

a) $f(x, y, z) = x^4 + y^4 + z^4$;

b) $f(x, y, z) = xyz$;

c) $f(x, y, z) = x^2 y^3 z^4$;

d) $f(x, y, z) = \text{sen } \sqrt{x^2 + y^2 + z^2}$.

13.1.2 Cálculo diferencial vectorial.

Ahora tenemos a mano estas operaciones básicas: gradiente, divergencia, rotacional y operador de Laplace. En la tabla siguiente se resumen sus propiedades y las relaciones entre ellas. Estas son fórmulas generales básicas, útiles cuando se trabaja con campos vectoriales en \mathbb{R}^3 .

En lo que sigue, f y g denotan campos escalares; \vec{F} , \vec{G} y \vec{H} denotan campos vectoriales.

Algunas expresiones en esta tabla requieren explicación.

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Primero, en la identidad 7, $\vec{V} = (\vec{F} \cdot \vec{\nabla}) \vec{G}$ tiene, por definición, componentes $V_i = \vec{F} \cdot (\vec{\nabla} G_i)$, para $i = 1, 2, 3$, donde $\vec{G} = (G_1, G_2, G_3)$.

Segundo, en la identidad 13, $\nabla^2 \vec{F}$ tiene componentes $\nabla^2 F_i$, donde $\vec{F} = (F_1, F_2, F_3)$.

1. $\vec{\nabla}(f + g) = \vec{\nabla}f + \vec{\nabla}g$
2. $\vec{\nabla}(cf) = c\vec{\nabla}f, \quad c \text{ const}$
3. $\vec{\nabla}(fg) = f\vec{\nabla}g + g\vec{\nabla}f$
4. $\vec{\nabla}(f/g) = (g\vec{\nabla}f - f\vec{\nabla}g)/g^2, \quad g(x) \neq 0$
5. $\text{div}(\vec{F} + \vec{G}) = \text{div}\vec{F} + \text{div}\vec{G}$
6. $\text{rot}(\vec{F} + \vec{G}) = \text{rot}\vec{F} + \text{rot}\vec{G}$
7. $\vec{\nabla}(\vec{F} \cdot \vec{G}) = (\vec{F} \cdot \vec{\nabla})\vec{G} + (\vec{G} \cdot \vec{\nabla})\vec{F} + \vec{F} \times \text{rot}\vec{G} + \vec{G} \times \text{rot}\vec{F}$
8. $\text{div}(f\vec{F}) = f\text{div}\vec{F} + \vec{F} \cdot \vec{\nabla}f$
9. $\text{div}(\vec{F} \times \vec{G}) = \vec{G} \cdot \text{rot}\vec{F} - \vec{F} \cdot \text{rot}\vec{G}$
10. $\text{div}\text{rot}\vec{F} = 0$
11. $\text{rot}(f\vec{F}) = f\text{rot}\vec{F} + \vec{\nabla}f \times \vec{F}$
12. $\text{rot}(\vec{F} \times \vec{G}) = \vec{F}\text{div}\vec{G} - \vec{G}\text{div}\vec{F} + (\vec{G} \cdot \vec{\nabla})\vec{F} - (\vec{F} \cdot \vec{\nabla})\vec{G}$
13. $\text{rot}\text{rot}\vec{F} = \text{grad}\text{div}\vec{F} - \nabla^2\vec{F}$
14. $\text{rot}\vec{\nabla}f = \vec{0}$
15. $\vec{\nabla}(\vec{F} \cdot \vec{F}) = 2(\vec{F} \cdot \vec{\nabla})\vec{F} + 2\vec{F} \times (\text{rot}\vec{F})$
16. $\nabla^2(fg) = f\nabla^2g + g\nabla^2f + 2(\vec{\nabla}f \cdot \vec{\nabla}g)$
17. $\text{div}(\vec{\nabla}f \times \vec{\nabla}g) = 0$
18. $\vec{\nabla} \cdot (f\vec{\nabla}g - g\vec{\nabla}f) = f\nabla^2g - g\nabla^2f$
19. $\vec{H} \cdot (\vec{F} \times \vec{G}) = \vec{G} \cdot (\vec{H} \times \vec{F}) = \vec{F} \cdot (\vec{G} \times \vec{H})$
20. $\vec{H} \cdot ((\vec{F} \times \vec{\nabla}) \times \vec{G}) = ((\vec{H} \cdot \vec{\nabla})\vec{G}) \cdot \vec{F} - (\vec{H} \cdot \vec{F}) (\vec{\nabla} \cdot \vec{G})$
21. $\vec{F} \times (\vec{G} \times \vec{H}) = (\vec{F} \cdot \vec{H})\vec{G} - \vec{H}(\vec{F} \cdot \vec{G})$

En la identidad 20, la expresión $(\vec{F} \times \vec{\nabla}) \times \vec{G}$, significa que $\vec{\nabla}$ opera sólo en \vec{G} .

Si lo comprobamos en detalle tendremos lo siguiente: empezando con la parte izquierda de la ecuación, hallando los términos por separado tenemos

$$\vec{F} \times \vec{\nabla} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ F_1 & F_2 & F_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{vmatrix} = (F_2 \frac{\partial}{\partial z} - F_3 \frac{\partial}{\partial y})\vec{i} - (F_1 \frac{\partial}{\partial z} - F_3 \frac{\partial}{\partial x})\vec{j} + (F_1 \frac{\partial}{\partial y} - F_2 \frac{\partial}{\partial x})\vec{k}$$

luego

$$\begin{aligned} (\vec{F} \times \vec{\nabla}) \times \vec{G} &= \begin{vmatrix} & \vec{i} & \vec{j} & \vec{k} \\ F_2 \frac{\partial}{\partial z} - F_3 \frac{\partial}{\partial y} & -F_1 \frac{\partial}{\partial z} + F_3 \frac{\partial}{\partial x} & F_1 \frac{\partial}{\partial y} - F_2 \frac{\partial}{\partial x} \\ G_1 & G_2 & G_3 \end{vmatrix} = \\ &= (-F_1 \frac{\partial G_3}{\partial z} + F_3 \frac{\partial G_3}{\partial x} - F_1 \frac{\partial G_2}{\partial y} + F_2 \frac{\partial G_2}{\partial x})\vec{i} - (F_2 \frac{\partial G_3}{\partial z} - F_3 \frac{\partial G_3}{\partial y} - F_1 \frac{\partial G_1}{\partial y} + F_2 \frac{\partial G_1}{\partial x})\vec{j} + \end{aligned}$$

13.1. CAMPOS VECTORIALES.

$$\begin{aligned}
 & +(F_2 \frac{\partial G_2}{\partial z} - F_3 \frac{\partial G_2}{\partial y} + F_1 \frac{\partial G_1}{\partial z} - F_3 \frac{\partial G_1}{\partial x}) \vec{k} \\
 \vec{H} \cdot ((\vec{F} \times \vec{\nabla}) \times \vec{G}) &= -F_1 \frac{\partial G_3}{\partial z} H_1 + F_3 \frac{\partial G_3}{\partial x} H_1 - F_1 \frac{\partial G_2}{\partial y} H_1 + F_2 \frac{\partial G_2}{\partial x} H_1 + \\
 & -F_2 \frac{\partial G_3}{\partial z} H_2 + F_3 \frac{\partial G_3}{\partial y} H_2 + F_1 \frac{\partial G_1}{\partial y} H_2 - F_2 \frac{\partial G_1}{\partial x} H_2 + \\
 & +F_2 \frac{\partial G_2}{\partial z} H_3 - F_3 \frac{\partial G_2}{\partial y} H_3 + F_1 \frac{\partial G_1}{\partial z} H_3 - F_3 \frac{\partial G_1}{\partial x} H_3
 \end{aligned}$$

Pasando a la parte derecha, por el primer término

$$\begin{aligned}
 (\vec{H} \cdot \vec{\nabla}) \vec{G} &= \left((H_1, H_2, H_3) \cdot \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \right) (G_1, G_2, G_3) = \\
 &= \left(H_1 \frac{\partial}{\partial x} + H_2 \frac{\partial}{\partial y} + H_3 \frac{\partial}{\partial z} \right) (G_1, G_2, G_3) = \left[H_1 \frac{\partial G_1}{\partial x} + H_2 \frac{\partial G_1}{\partial y} + H_3 \frac{\partial G_1}{\partial z} \right] \vec{i} + \\
 &+ \left[H_1 \frac{\partial G_2}{\partial x} + H_2 \frac{\partial G_2}{\partial y} + H_3 \frac{\partial G_2}{\partial z} \right] \vec{j} + \left[H_1 \frac{\partial G_3}{\partial x} + H_2 \frac{\partial G_3}{\partial y} + H_3 \frac{\partial G_3}{\partial z} \right] \vec{k} \\
 ((\vec{H} \cdot \vec{\nabla}) \vec{G}) \cdot \vec{F} &= F_1 \frac{\partial G_1}{\partial x} H_1 + F_1 \frac{\partial G_1}{\partial y} H_2 + F_1 \frac{\partial G_1}{\partial z} H_3 + F_2 \frac{\partial G_2}{\partial x} H_1 + F_2 \frac{\partial G_2}{\partial y} H_2 + \\
 &+ F_2 \frac{\partial G_2}{\partial z} H_3 + F_3 \frac{\partial G_3}{\partial x} H_1 + F_3 \frac{\partial G_3}{\partial y} H_2 + F_3 \frac{\partial G_3}{\partial z} H_3
 \end{aligned}$$

Además, por el segundo término

$$\begin{aligned}
 (\vec{H} \cdot \vec{F}) (\vec{\nabla} \cdot \vec{G}) &= [(H_1, H_2, H_3) \cdot (F_1, F_2, F_3)] \left[\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot (G_1, G_2, G_3) \right] = \\
 &= (H_1 F_1 + H_2 F_2 + H_3 F_3) \left(\frac{\partial G_1}{\partial x} + \frac{\partial G_2}{\partial y} + \frac{\partial G_3}{\partial z} \right) = \\
 &= F_1 \frac{\partial G_1}{\partial x} H_1 + F_1 \frac{\partial G_2}{\partial y} H_1 + F_1 \frac{\partial G_3}{\partial z} H_1 + F_2 \frac{\partial G_1}{\partial x} H_2 + F_2 \frac{\partial G_2}{\partial y} H_2 + \\
 &+ F_2 \frac{\partial G_3}{\partial z} H_2 + F_3 \frac{\partial G_1}{\partial x} H_3 + F_3 \frac{\partial G_2}{\partial y} H_3 + F_3 \frac{\partial G_3}{\partial z} H_3
 \end{aligned}$$

Restando estos dos últimos términos:

$$\begin{aligned}
 & ((\vec{H} \cdot \vec{\nabla}) \vec{G}) \cdot \vec{F} - (\vec{H} \cdot \vec{F}) (\vec{\nabla} \cdot \vec{G}) = \\
 &= F_1 \frac{\partial G_1}{\partial x} H_1 + F_1 \frac{\partial G_1}{\partial y} H_2 + F_1 \frac{\partial G_1}{\partial z} H_3 + F_2 \frac{\partial G_2}{\partial x} H_1 + F_2 \frac{\partial G_2}{\partial y} H_2 + F_2 \frac{\partial G_2}{\partial z} H_3 + \\
 &+ F_3 \frac{\partial G_3}{\partial x} H_1 + F_3 \frac{\partial G_3}{\partial y} H_2 + F_3 \frac{\partial G_3}{\partial z} H_3 - F_1 \frac{\partial G_1}{\partial x} H_1 - F_1 \frac{\partial G_2}{\partial y} H_1 - F_1 \frac{\partial G_3}{\partial z} H_1 + \\
 &- F_2 \frac{\partial G_1}{\partial x} H_2 - F_2 \frac{\partial G_2}{\partial y} H_2 - F_2 \frac{\partial G_3}{\partial z} H_2 - F_3 \frac{\partial G_1}{\partial x} H_3 - F_3 \frac{\partial G_2}{\partial y} H_3 - F_3 \frac{\partial G_3}{\partial z} H_3
 \end{aligned}$$

que coincide con la expresión obtenida anteriormente para la parte izquierda de la igualdad.

Dejamos la comprobación de las demás identidades como ejercicio.

Ejemplo. Sea \vec{r} el campo vectorial $\vec{r}(x, y, z) = (x, y, z)$, es decir el vector posición, y sea $r = \|\vec{r}\|$. Calcular

$$\text{grad } r = \vec{\nabla} r \quad \text{y} \quad \text{div } (r \vec{r}) = \vec{\nabla} \cdot (r \vec{r}).$$

Solución:

Dado que $\vec{r}(x, y, z) = (x, y, z)$, se sigue que $r = \|\vec{r}\| = \sqrt{x^2 + y^2 + z^2}$. Entonces

$$\text{grad } r = \vec{\nabla} r = \left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z} \right) = \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r} \right) = \frac{\vec{r}}{r}.$$

Para la segunda parte, usaremos la identidad 8, y el hecho que

$$\text{div } \vec{r} = \vec{\nabla} \cdot \vec{r} = \frac{\partial r}{\partial x} + \frac{\partial r}{\partial y} + \frac{\partial r}{\partial z} = 3 \quad \text{y} \quad \vec{r} \cdot \vec{\nabla} r = \vec{r} \cdot \frac{\vec{r}}{r} = \frac{r^2}{r} = r,$$

para escribir

$$\text{div } (r \vec{r}) = \vec{\nabla} \cdot (r \vec{r}) = r \text{div } \vec{r} + \vec{r} \cdot \vec{\nabla} r = 3r + r = 4r.$$

Ejercicios.

- 1) Sea f una función diferenciable de una variable y sea el campo vectorial $\vec{v}(x, y, z)$ dado por $\vec{v}(x, y, z) = f(x)\vec{i}$. Determina la divergencia $\vec{\nabla} \cdot \vec{v}$ y el rotacional $\vec{\nabla} \times \vec{v}$.
- 2) Demuestra que si $\vec{v}(\vec{r})$ es un campo vectorial diferenciable de la forma $\vec{v}(\vec{r}) = f(x)\vec{i} + g(y)\vec{j} + h(z)\vec{k}$, se verifica que $\vec{\nabla} \times \vec{v} = \vec{0}$.

13.2 La integral de trayectoria.

En esta sección se introduce el concepto de integral de trayectoria; ésta es una de las maneras en que se pueden generalizar las integrales de funciones de una variable, a funciones de varias variables. En el capítulo 12 estudiamos las integrales dobles y triples; en esta y otras secciones posteriores estudiaremos más generalizaciones.

Supongamos que $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ es una función escalar definida en el espacio tridimensional. Será útil definir la integral de tal función a lo largo de la trayectoria $\vec{\sigma}: U = [a, b] \rightarrow \mathbb{R}^3$, con $\vec{\sigma}(t) = (x(t), y(t), z(t))$.

Definición. La **integral de trayectoria**, o **integral de $f(x, y, z)$ a lo largo de la trayectoria $\vec{\sigma}$** , está definida cuando $\vec{\sigma}: U = [a, b] \rightarrow \mathbb{R}^3$ es de clase \mathcal{C}^1 y cuando la función compuesta $f \circ \vec{\sigma}: t \rightsquigarrow f(x(t), y(t), z(t))$ es continua en I . Definimos esta integral por la ecuación

$$\int_{\vec{\sigma}} f \, ds = \int_a^b f(x(t), y(t), z(t)) \|\vec{\sigma}'(t)\| \, dt = \int_a^b f(\vec{\sigma}(t)) \|\vec{\sigma}'(t)\| \, dt. \quad (13.8)$$

13.2. LA INTEGRAL DE TRAYECTORIA.

Nótese que basta que la función f esté definida en la curva imagen de $\vec{\sigma}$, y no necesariamente en todo el espacio, para que tenga sentido la definición dada.

Si $\vec{\sigma}(t)$ sólo es C^1 a trozos o $f(\vec{\sigma}(t))$ es continua a trozos, definimos la integral de trayectoria rompiendo el intervalo $[a, b]$ en partes sobre las cuales $f(\vec{\sigma}(t)) \|\vec{\sigma}'(t)\|$ sea continua, y sumando las integrales sobre las partes.

Finalmente, si $f = 1$, estamos simplemente reenunciando la definición de longitud de arco de $\vec{\sigma}$, que dimos en la fórmula (10.3).

Ejemplo. Sea $\vec{\sigma}$ la hélice $\vec{\sigma}: [0, 2\pi] \rightarrow \mathbb{R}^3$, dada por $t \rightsquigarrow (\cos t, \sin t, t)$, y sea la función escalar definida por $f(x, y, z) = x^2 + y^2 + z^2$. Evalúa la integral de trayectoria

$$\int_{\vec{\sigma}} f ds.$$

Solución:

Tenemos

$$\|\vec{\sigma}'(t)\| = \sqrt{\left[\frac{d \cos t}{dt}\right]^2 + \left[\frac{d \sin t}{dt}\right]^2 + \left[\frac{dt}{dt}\right]^2} = \sqrt{[-\sin t]^2 + [\cos t]^2 + [1]^2} = \sqrt{2}.$$

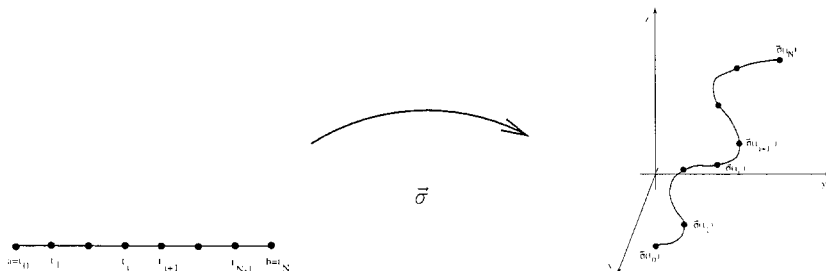
Sustituimos las expresiones de x , y y z en la función f para obtener, a lo largo de $\vec{\sigma}$

$$f(x(t), y(t), z(t)) = f(\vec{\sigma}(t)) = \cos^2 t + \sin^2 t + t^2 = 1 + t^2$$

Luego, usando la fórmula (13.8)

$$\int_{\vec{\sigma}} f ds = \int_0^{2\pi} f(\vec{\sigma}(t)) \|\vec{\sigma}'(t)\| dt = \int_0^{2\pi} (1 + t^2) \sqrt{2} dt = \frac{2\sqrt{2}\pi}{3} (3 + 4\pi^2).$$

Para motivar la definición de integral de trayectoria, consideraremos sumas S_N del tipo Riemann. Para simplificar, supongamos que $\vec{\sigma}$ es de clase C^1 en I . Podemos subdividir el intervalo $I = [a, b]$ por medio de una partición $a = t_0 < t_1 < \dots < t_N = b$. Esto produce una descomposición de $\vec{\sigma}$ en trayectorias $\vec{\sigma}_i$ definidas en $[t_i, t_{i+1}]$, para $0 \leq i \leq N - 1$.



Si denotamos la longitud de arco de $\vec{\sigma}_i$ por Δs_i , tenemos

$$\Delta s_i = \int_{t_i}^{t_{i+1}} \|\vec{\sigma}'(t)\| dt.$$

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Cuando N es grande, la longitud de arco Δs_i es pequeña y $f(x, y, z)$ es aproximadamente constante para puntos en $\vec{\sigma}_i$. Si consideramos las sumas

$$S_N = \sum_{i=0}^{N-1} f(x_i, y_i, z_i) \Delta s_i$$

donde $(x_i, y_i, z_i) = \vec{\sigma}(t)$ para algún $t \in [t_i, t_{i+1}]$, estas sumas son básicamente sumas de Riemann, luego

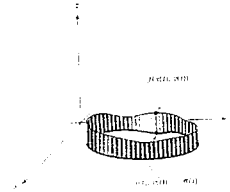
$$\lim_{N \rightarrow \infty} S_N = \int_I f(x(t), y(t), z(t)) \|\vec{\sigma}'(t)\| dt = \int_{\vec{\sigma}} f ds.$$

Así, la integral de trayectoria se puede expresar como límite de sumas de Riemann.

Un caso particular e importante de la integral de trayectoria se presenta cuando dicha trayectoria $\vec{\sigma}$ describe una curva plana. Si todos los puntos $\vec{\sigma}(t)$ están en el plano xy y si f es una función de dos variables reales con valores reales, la integral de trayectoria de f a lo largo de $\vec{\sigma}$ es

$$\int_{\vec{\sigma}} f(x, y) ds = \int_a^b f(x(t), y(t)) \|\vec{\sigma}'(t)\| dt = \int_a^b f(x(t), y(t)) \sqrt{[x'(t)]^2 + [y'(t)]^2} dt.$$

Cuando $f(x, y) \geq 0$, esta integral tiene una interpretación geométrica como el "área de una valla". Podemos construir una valla cuya base sea la imagen de $\vec{\sigma}$ y altura $f(x, y)$ en (x, y) . Si $\vec{\sigma}$ recorre sólo una vez la imagen de $\vec{\sigma}$, la integral de trayectoria representa el área de un lado de la valla.



Ejercicios.

1) Evalúa las siguientes integrales de trayectorias $\int_{\vec{\sigma}} f(x, y, z) ds$, donde

- a) $f(x, y, z) = x + y + z$ y $\vec{\sigma}: t \rightsquigarrow (\text{sen } t, \cos t, t)$, para $t \in [0, 2\pi]$;
- b) $f(x, y, z) = \cos z$ y $\vec{\sigma}: t \rightsquigarrow (\text{sen } t, \cos t, t)$, para $t \in [0, 2\pi]$;
- c) $f(x, y, z) = y$ y $\vec{\sigma}: t \rightsquigarrow (0, 0, t)$, para $t \in [0, 1]$;
- d) $f(x, y, z) = x \cos z$ y $\vec{\sigma}: t \rightsquigarrow (t, t^2, 0)$, para $t \in [0, 1]$;
- e) $f(x, y, z) = yz$ y $\vec{\sigma}: t \rightsquigarrow (t, 3t, 2t)$, para $t \in [1, 3]$.

2) Halla $l(\vec{\sigma})$, la longitud de la trayectoria dada por $\vec{\sigma}(t) = (t^2, t, 3)$, para $t \in [0, 1]$.

13.3 La integral de línea.

La noción de trabajo nos llevará a la definición de la **integral de línea**. El trabajo realizado por una fuerza constante \vec{F} que actúa sobre un punto material que se mueve a lo largo de una recta es, por definición, el componente de \vec{F} en la dirección del desplazamiento multiplicado por la longitud del vector \vec{d} determinado por el desplazamiento:

$$W = (\text{comp}_d \vec{F}) \cdot \|\vec{d}\| = \vec{F} \cdot \vec{d}.$$

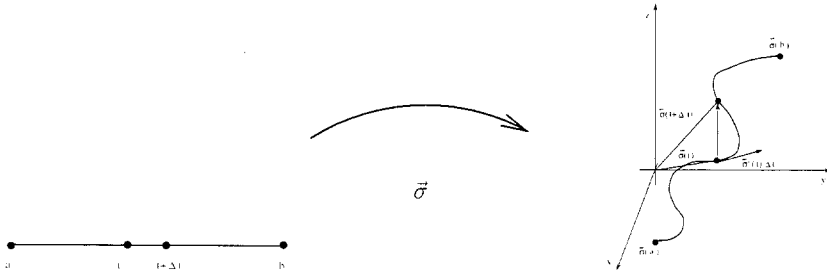
Esta noción elemental de trabajo es útil pero no es suficiente. Consideremos, por ejemplo, el caso de un punto material que se mueve en un campo eléctrico o en un campo gravitatorio. En general, la trayectoria del punto no es recta sino una curva, y la fuerza, en lugar de ser constante, tiende a variar de un punto a otro. Lo que queremos es una noción de trabajo que se aplique en estas situaciones más generales.

Como en el caso de la deducción de la fórmula (10.3) para la longitud de arco y la (13.8) para la integral de trayectoria, se puede justificar la deducción de la siguiente fórmula para el trabajo realizado por el campo de fuerza \vec{F} sobre una partícula que se mueve a lo largo de la trayectoria $\vec{\sigma}: [a, b] \rightarrow \mathbb{R}^3$:

$$\text{trabajo realizado por } \vec{F} = \int_a^b \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt. \quad (13.9)$$

Conforme t varía sobre un pequeño intervalo $t + \Delta t$, la partícula se mueve de $\vec{\sigma}(t)$ a $\vec{\sigma}(t + \Delta t)$, con un vector de desplazamiento dado por $\Delta \vec{s} = \vec{\sigma}(t + \Delta t) - \vec{\sigma}(t) \approx \vec{\sigma}'(t) \Delta t$, (esta última aproximación se obtiene de la definición de derivada). El trabajo realizado al ir de $\vec{\sigma}(t)$ a $\vec{\sigma}(t + \Delta t)$ es, por lo tanto, aproximadamente

$$\vec{F}(\vec{\sigma}(t)) \cdot \Delta \vec{s} \approx \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) \Delta t.$$



Si subdividimos el intervalo $[a, b]$ en N partes iguales, $a = t_0 < t_1 < \dots < t_N = b$, con $\Delta t = t_{i+1} - t_i = \frac{b-a}{N}$, entonces el trabajo realizado por \vec{F} es aproximadamente

$$\sum_{i=0}^{n-1} \vec{F}(\vec{\sigma}(t_i)) \cdot \Delta \vec{s}_i \approx \sum_{i=0}^{n-1} \vec{F}(\vec{\sigma}(t_i)) \cdot \vec{\sigma}'(t_i) \Delta t.$$

Cuando $n \rightarrow \infty$, esta aproximación se vuelve cada vez mejor, de modo que es razonable definir el trabajo como el límite de tal suma cuando $n \rightarrow \infty$. Pero este límite está dado por la integral $\int_a^b \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt$. Llegamos entonces a la siguiente definición.

Definición. Sea \vec{F} un campo vectorial en \mathbb{R}^3 continuo sobre la trayectoria $\vec{\sigma}(t)$ de clase C^1 , $\vec{\sigma}: [a, b] \rightarrow \mathbb{R}^3$. La **integral de línea de \vec{F} a lo largo de $\vec{\sigma}$** viene dada por la fórmula

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_a^b \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt. \quad (13.10)$$

Si una masa puntual m se mueve de tal forma que en el instante t su posición es $\vec{r}(t)$, según la segunda ley de Newton, la fuerza total que actúa sobre dicha masa puntual en el instante t ha de ser $m\vec{r}''(t)$.

Ejemplo 1. Una masa puntual m se mueve desde $t = 0$ hasta $t = 1$ de tal forma que en el instante t su posición viene dada por

$$\vec{r}(t) = \alpha t^2 \vec{i} + \text{sen } \beta t \vec{j} + \cos \beta t \vec{k}.$$

Halla la fuerza total que actúa sobre esta masa puntual en el instante t y calcula el trabajo total realizado por dicha fuerza.

Solución:

La diferenciación del campo vectorial $\vec{r}(t)$ nos da

$$\vec{r}'(t) = 2\alpha t \vec{i} + \beta \cos \beta t \vec{j} - \beta \text{sen } \beta t \vec{k} \quad \vec{r}''(t) = 2\alpha \vec{i} - \beta^2 \text{sen } \beta t \vec{j} - \beta^2 \cos \beta t \vec{k}.$$

Luego la fuerza total sobre el objeto en el instante t es

$$\vec{F}(t) = m\vec{r}''(t) = m \left(2\alpha \vec{i} - \beta^2 \text{sen } \beta t \vec{j} - \beta^2 \cos \beta t \vec{k} \right).$$

Podemos calcular el trabajo realizado por esta fuerza integrando la fuerza sobre la curva

$$\vec{\sigma}: \quad \vec{\sigma}(t) = \vec{r}(t) = \alpha t^2 \vec{i} + \text{sen } \beta t \vec{j} + \cos \beta t \vec{k}, \quad t \in [0, 1].$$

Esto nos da

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_0^1 \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt = \int_0^1 m\vec{r}''(t) \cdot \vec{r}'(t) dt = m \int_0^1 4\alpha^2 t dt = 2\alpha^2 m.$$

Ejemplo 2. Sea $\vec{\sigma}(t) = (\text{sen } t, \cos t, t)$ con $0 \leq t \leq 2\pi$. Sea $\vec{F}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$. Calcula la integral de línea

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}.$$

13.3. LA INTEGRAL DE LÍNEA.

Solución:

Aquí,

$$\vec{F}(\vec{\sigma}(t)) = \vec{F}(\sin t, \cos t, t) = \sin t \vec{i} + \cos t \vec{j} + t \vec{k}$$

y

$$\vec{\sigma}'(t) = \cos t \vec{i} - \sin t \vec{j} + \vec{k},$$

por lo tanto

$$\vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) = \sin t \cos t - \cos t \sin t + t = t$$

y entonces, usando la fórmula (13.10)

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_0^{2\pi} \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt = \int_0^{2\pi} t dt = 2\pi^2.$$

Ejemplo 3. Integra el campo vectorial $\vec{h}(x, y, z) = xy\vec{i} + yz\vec{j} + xz\vec{k}$ sobre la cúbica alabeada $\vec{\sigma}(t) = t\vec{i} + t^2\vec{j} + t^3\vec{k}$, desde $(-1, 1, -1)$ hasta $(1, 1, 1)$.

Solución:

La curva para la integración empieza en $t = -1$ y termina en $t = 1$. En este caso es

$$x(t) = t, \quad y(t) = t^2, \quad z(t) = t^3 \quad \Rightarrow \quad x'(t) = 1, \quad y'(t) = 2t, \quad z'(t) = 3t^2.$$

Luego

$$\begin{aligned} \vec{h}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) &= \left(x(t)y(t)\vec{i} + y(t)z(t)\vec{j} + x(t)z(t)\vec{k} \right) \cdot \left(x'(t)\vec{i} + y'(t)\vec{j} + z'(t)\vec{k} \right) = \\ &= x(t)y(t)x'(t) + y(t)z(t)y'(t) + x(t)z(t)z'(t) = \\ &= t(t^2)1 + t^2(t^3)(2t) + t(t^3)(3t^2) = t^3 + 5t^6. \end{aligned}$$

Finalmente, usando la fórmula (13.10)

$$\int_{\vec{\sigma}} \vec{h} \cdot d\vec{s} = \int_0^{2\pi} \vec{h}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt = \int_{-1}^1 (t^3 + 5t^6) dt = 2 \int_0^1 5t^6 dt = \frac{10}{7}.$$

Hay otra fórmula útil para la integral de línea en el caso de trayectorias $\vec{\sigma}$ que satisfagan

$\vec{\sigma}'(t) \neq \vec{0}$: si $\vec{T}(t) = \frac{\vec{\sigma}'(t)}{\|\vec{\sigma}'(t)\|}$ denota el vector tangente unitario, tenemos

$$\begin{aligned} \int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} &= \int_a^b \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt = \int_a^b \left[\vec{F}(\vec{\sigma}(t)) \cdot \frac{\vec{\sigma}'(t)}{\|\vec{\sigma}'(t)\|} \right] \|\vec{\sigma}'(t)\| dt = \\ &= \int_a^b \left[\vec{F}(\vec{\sigma}(t)) \cdot \vec{T}(t) \right] \|\vec{\sigma}'(t)\| dt. \end{aligned} \tag{13.11}$$

Escrita de esta manera, la fórmula (13.11) nos dice que la integral de línea $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}$ es igual a la integral de trayectoria de la componente tangencial $\vec{F}(\vec{\sigma}(t)) \cdot \vec{T}(t)$ de \vec{F} a lo largo de $\vec{\sigma}$.

Para calcular una integral de línea en cualquier caso particular, podemos usar la definición original (13.10) o podemos integrar la componente tangencial de \vec{F} a lo largo de $\vec{\sigma}$, usando la fórmula (13.11), dependiendo de que sea más fácil o más apropiado.

Otra forma de escribir la integral de línea es

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_{\vec{\sigma}} F_1 dx + F_2 dy + F_3 dz. \quad (13.12)$$

donde F_1, F_2, F_3 son las componentes del campo vectorial \vec{F} . A la expresión que aparece como integrando, $F_1 dx + F_2 dy + F_3 dz$ la llamamos **forma diferencial**. Por definición la integral de una forma diferencial es

$$\int_{\vec{\sigma}} F_1 dx + F_2 dy + F_3 dz = \int_{\vec{\sigma}} \left(F_1 \frac{dx}{dt} + F_2 \frac{dy}{dt} + F_3 \frac{dz}{dt} \right) dt = \int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}.$$

Ejemplo. Evalúa

$$\int_{\vec{\sigma}} x^2 dx + xy dy + dz$$

si la trayectoria $\vec{\sigma}: [0, 1] \rightarrow \mathbb{R}^3$ está dada por $\vec{\sigma}(t) = (t, t^2, 1)$.

Solución:

Dado que

$$\frac{dx}{dt} = 1, \quad \frac{dy}{dt} = 2t, \quad \frac{dz}{dt} = 0$$

se sigue que

$$\begin{aligned} \int_{\vec{\sigma}} x^2 dx + xy dy + dz &= \int_{\vec{\sigma}} \left(x^2(t) \frac{dx}{dt} + x(t)y(t) \frac{dy}{dt} + \frac{dz}{dt} \right) dt = \\ &= \int_0^1 (t^2 + t \cdot t^2 \cdot (2t)) dt = \int_0^1 (t^2 + 2t^4) dt = \frac{11}{15}. \end{aligned}$$

Como sucede para las funciones escalares también podemos definir la integral de línea (13.10) si el integrando $\vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t)$ sólo es continuo a trozos. Análogamente, si una curva no es suave pero está formada por la unión de un número finito de trozos suaves contiguos, definimos la integral sobre la curva como la suma sobre los trozos. Diremos que una curva de estas características es **suave a trozos**. Todos los polígono son suaves a trozos.

Definición. Definimos una **curva simple** C como la imagen de una función C^1 a trozos, $\vec{\sigma}: I \rightarrow \mathbb{R}^3$ y que sea uno a uno en un intervalo I ; $\vec{\sigma}$ se llama **parametrización** de C . Así, una curva simple es aquella que no se intersecta a sí misma. Si $I = [a, b]$, llamamos $\vec{\sigma}(a)$ y $\vec{\sigma}(b)$ **extremos** de la curva. Cada curva simple C tiene dos orientaciones o direcciones asociadas con ella. Si P y Q son los extremos de la curva, entonces podemos considerar C como dirigida ya sea de P a Q o de Q a P . La curva simple C junto con un sentido de dirección se llama **curva simple orientada** o curva simple dirigida.

13.3. LA INTEGRAL DE LÍNEA.

Definición. Por **curva cerrada simple** entenderemos la imagen de una función C^1 a trozos, $\vec{\sigma}: I \rightarrow \mathbb{R}^3$, que sea uno a uno en un intervalo $[a, b]$ y que satisfaga la condición $\vec{\sigma}(a) = \vec{\sigma}(b)$. Las curvas cerradas simples tienen dos orientaciones, correspondientes a las dos direcciones de movimiento posibles a lo largo de la curva, en el sentido de las agujas de un reloj o en sentido contrario a ellas.

Si C es una curva simple orientada o una curva cerrada simple orientada, podemos definir sin ambigüedad la integral de trayectoria

$$\int_C f \, ds = \int_{\vec{\sigma}} f \, ds,$$

y la integral de línea

$$\int_C \vec{F} \cdot d\vec{s} = \int_{\vec{\sigma}} \vec{F} \cdot d\vec{s},$$

donde $\vec{\sigma}$ es cualquier parametrización uno a uno que preserve la orientación de C .

Una curva cerrada se puede parametrizar de varias maneras, y tendremos que tener cuidado porque es posible que dos funciones $\vec{\sigma}$ y $\vec{\eta}$ tengan la misma imagen, e induzcan la misma orientación en la imagen, pero las integrales de línea del mismo campo vectorial no coincidan. Veamos este hecho en el siguiente ejemplo.

Ejemplo 1. Sean $\vec{\sigma}$ y $\vec{\eta}$ dos parametrizaciones de círculo unitario C en el plano xy , que recorren el círculo en la misma dirección

$$\vec{\sigma}(t) = (\cos t, \sin t, 0), \quad \text{para } 0 \leq t \leq 2\pi.$$

y

$$\vec{\eta}(t) = (\cos(2t), \sin(2t), 0), \quad \text{para } 0 \leq t \leq 2\pi.$$

Calcula

$$\int_C \vec{F} \cdot d\vec{s}$$

si el campo vectorial es $\vec{F}(x, y, z) = (y, 0, 0)$ y usando las dos parametrizaciones dadas.

Solución:

Con la primera parametrización $\vec{\sigma}(t) = (\cos t, \sin t, 0)$ obtenemos

$$\int_C \vec{F} \cdot d\vec{s} = \int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_0^{2\pi} F_1(\vec{\sigma}(t)) \frac{dx}{dt} dt = - \int_0^{2\pi} \sin^2 t \, dt = -\pi.$$

y con la segunda parametrización $\vec{\eta}(t) = (\cos(2t), \sin(2t), 0)$ obtenemos

$$\int_C \vec{F} \cdot d\vec{s} = \int_{\vec{\eta}} \vec{F} \cdot d\vec{s} = \int_0^{2\pi} F_1(\vec{\eta}(t)) \frac{dx}{dt} dt = -2 \int_0^{2\pi} \sin^2(2t) \, dt = -2\pi.$$

La razón para esto es que la primera parametrización $\vec{\sigma}$ es uno a uno pero la segunda $\vec{\eta}$ no lo es: $\vec{\eta}$ recorre el círculo unitario dos veces en dirección contraria a la que giran las

manecillas del reloj; por lo tanto $\vec{\eta}$ no es una parametrización del círculo unitario, como una curva cerrada simple.

En el problema siguiente integramos sobre un triángulo: lo hacemos integrando sobre cada lado y sumando después los resultados.

Ejemplo 2. Calcula la integral

$$\int_C \vec{F} \cdot d\vec{s}$$

si $\vec{F}(x, y) = e^y \vec{i} - \text{sen}(\pi x) \vec{j}$ y C representa el triángulo con vértices $(1, 0)$, $(0, 1)$, $(-1, 0)$ recorrido en sentido contrario al de las agujas de un reloj.

Solución:

El triángulo está formado por tres segmentos de recta, que podemos parametrizar según

$$C_1: \vec{\sigma}_1(t) = (1-t)\vec{i} + t\vec{j}, \quad t \in [0, 1],$$

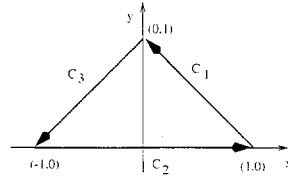
$$x_1(t) = 1-t, \quad y_1(t) = t;$$

$$C_2: \vec{\sigma}_2(t) = -t\vec{i} + (1-t)\vec{j}, \quad t \in [0, 1],$$

$$x_2(t) = -t, \quad y_2(t) = 1-t;$$

$$C_3: \vec{\sigma}_3(t) = (2t-1)\vec{i}, \quad t \in [0, 1],$$

$$x_3(t) = 2t-1, \quad y_3(t) = 0;$$



por tanto tendremos

$$\begin{aligned} \int_{C_1} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_1} \vec{F}(\vec{\sigma}_1(t)) \cdot \vec{\sigma}'_1(t) dt = \int_0^1 [e^{y_1(t)} x'_1(t) - \text{sen}(\pi x_1(t)) y'_1(t)] dt = \\ &= \int_0^1 [-e^t - \text{sen}(\pi(1-t))] dt = 1 - e - \frac{2}{\pi}; \end{aligned}$$

$$\begin{aligned} \int_{C_2} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_2} \vec{F}(\vec{\sigma}_2(t)) \cdot \vec{\sigma}'_2(t) dt = \int_0^1 [e^{y_2(t)} x'_2(t) - \text{sen}(\pi x_2(t)) y'_2(t)] dt = \\ &= \int_0^1 [-e^{(1-t)} + \text{sen}(-\pi t)] dt = 1 - e - \frac{2}{\pi}; \end{aligned}$$

$$\begin{aligned} \int_{C_3} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_3} \vec{F}(\vec{\sigma}_3(t)) \cdot \vec{\sigma}'_3(t) dt = \int_0^1 [e^{y_3(t)} x'_3(t) - \text{sen}(\pi x_3(t)) y'_3(t)] dt = \\ &= \int_0^1 2 dt = 2. \end{aligned}$$

La integral sobre el triángulo es la suma de esas integrales:

$$\int_C \vec{F} \cdot d\vec{s} = \left(1 - e - \frac{2}{\pi}\right) + \left(1 - e - \frac{2}{\pi}\right) + 2 = 4 - 2e - \frac{4}{\pi} \approx -2.71.$$

En este último ejemplo vemos que el trabajo realizado por el campo es una cantidad negativa. Esto significa que el campo impide el movimiento a lo largo de la trayectoria.

13.3. LA INTEGRAL DE LÍNEA.

La integral de línea $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}$ depende no sólo del campo \vec{F} sino también de la trayectoria $\vec{\sigma}: [a, b] \rightarrow \mathbb{R}^3$. En general, si $\vec{\sigma}$ y $\vec{\rho}$ son dos trayectorias diferentes en \mathbb{R}^3 , tendremos $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} \neq \int_{\vec{\rho}} \vec{F} \cdot d\vec{s}$. Por otro lado, veremos que $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \pm \int_{\vec{\rho}} \vec{F} \cdot d\vec{s}$ se cumple para todo campo vectorial \vec{F} si $\vec{\rho}$ es lo que llamamos una reparametrización de $\vec{\sigma}$.

Definición. Sea $h: I \rightarrow I_1$ una función de clase C^1 con valores reales que sea una correspondencia biunívoca entre un intervalo $I = [a, b]$ sobre otro intervalo $I_1 = [a_1, b_1]$. Sea $\vec{\sigma}: I_1 \rightarrow \mathbb{R}^3$ una trayectoria C^1 a trozos. Entonces llamamos **reparametrización de $\vec{\sigma}$** a la composición

$$\vec{\rho} = \vec{\sigma} \circ h: I \rightarrow \mathbb{R}^3.$$

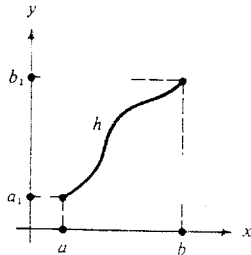
Esto significa que $\vec{\rho}(t) = \vec{\sigma}(h(t))$, de modo que se puede pensar en h como un cambio en la rapidez con que se mueve el punto a lo largo de la trayectoria. En efecto, resulta que $\vec{\rho}'(t) = \vec{\sigma}'(h(t)) h'(t)$, de modo que el vector velocidad para $\vec{\sigma}$ se multiplica por el factor escalar $h'(t)$.

Está implícito en la definición que h debe mandar extremos a extremos; esto es, se cumplirá

$$\vec{\sigma} \circ h(a) = \vec{\sigma}(a_1) \quad \text{y} \quad \vec{\sigma} \circ h(b) = \vec{\sigma}(b_1),$$

o

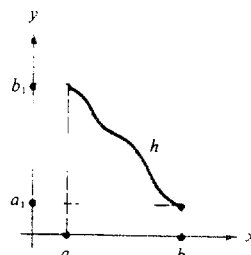
$$\vec{\sigma} \circ h(a) = \vec{\sigma}(b_1) \quad \text{y} \quad \vec{\sigma} \circ h(b) = \vec{\sigma}(a_1).$$



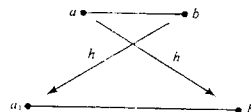
gráfica de h



h preserva la orientación



gráfica de h



h invierte la orientación

En el primer caso se dice que la reparametrización **preserva la orientación**, y una partícula que trace la trayectoria $\vec{\sigma} \circ h$ se mueve en la misma dirección que una partícula que trace la trayectoria $\vec{\sigma}$. En el segundo caso, la reparametrización **invierte la orientación**, y una partícula que trace la trayectoria $\vec{\sigma} \circ h$ se mueve en dirección opuesta a la de una partícula que trace la trayectoria $\vec{\sigma}$.

La trayectoria $\bar{\sigma}_{op}: [a, b] \rightarrow \mathbb{R}^3$, dada por $t \rightsquigarrow \bar{\sigma}(a + b - t)$, es una reparametrización de $\bar{\sigma}$ correspondiente a la función $h: [a, b] \rightarrow [a, b]$, dada por $t \rightsquigarrow a + b - t$. Llamamos $\bar{\sigma}_{op}$ a la **trayectoria opuesta** a $\bar{\sigma}$. Esta reparametrización invierte la orientación.

La trayectoria $\bar{\rho}: [0, 1] \rightarrow \mathbb{R}^3$, dada por $t \rightsquigarrow \bar{\sigma}(a + (b - a)t)$, es una reparametrización de $\bar{\sigma}$ que preserva la orientación, correspondiente al cambio de coordenadas $h: [0, 1] \rightarrow [a, b]$, dado por $t \rightsquigarrow a + (b - a)t$.

Veamos ahora un teorema, que nos permite concluir que si es conveniente reparametrizar una trayectoria cuando se evalúa una integral de línea, entonces el resultado de la integral no cambiará, excepto, quizá, el signo, dependiendo de la orientación.

Teorema 13.3 Sea \vec{F} un campo vectorial continuo en la trayectoria $C^1 \bar{\sigma}: [a_1, b_1] \rightarrow \mathbb{R}^3$ y sea $\bar{\rho}: [a, b] \rightarrow \mathbb{R}^3$ una reparametrización de $\bar{\sigma}$. Entonces

$$\int_{\bar{\sigma}} \vec{F} \cdot d\vec{s} = \begin{cases} \int_{\bar{\rho}} \vec{F} \cdot d\vec{s} & \text{si } \bar{\rho} \text{ preserva la orientación.} \\ - \int_{\bar{\rho}} \vec{F} \cdot d\vec{s} & \text{si } \bar{\rho} \text{ invierte la orientación.} \end{cases} \quad (13.13)$$

Demostración:

Por hipótesis, tenemos una función h tal que $\bar{\rho} = \bar{\sigma} \circ h$; esto es $\bar{\rho}(t) = \bar{\sigma}(h(t))$. Por la regla de la cadena,

$$\bar{\rho}'(t) = \bar{\sigma}'(h(t)) h'(t)$$

de modo que

$$\int_{\bar{\rho}} \vec{F} \cdot d\vec{s} = \int_a^b \vec{F}(\bar{\rho}(t)) \cdot \bar{\rho}'(t) dt = \int_a^b \vec{F}(\bar{\sigma}(h(t))) \cdot \bar{\sigma}'(h(t)) h'(t) dt.$$

Cambiando la variable con $s = h(t)$ (ver teorema 3.4 y la fórmula (3.2) o la (12.13)), obtenemos

$$\int_{h(a)}^{h(b)} \vec{F}(\bar{\sigma}(s)) \cdot \bar{\sigma}'(s) ds = \begin{cases} \int_{a_1}^{b_1} \vec{F}(\bar{\sigma}(s)) \cdot \bar{\sigma}'(s) ds = \int_{\bar{\sigma}} \vec{F} \cdot d\vec{s} \\ \int_{b_1}^{a_1} \vec{F}(\bar{\sigma}(s)) \cdot \bar{\sigma}'(s) ds = - \int_{\bar{\sigma}} \vec{F} \cdot d\vec{s} \end{cases}$$

respectivamente, si $\bar{\rho}$ preserva o invierte la orientación.

c.q.d

Ejemplo. Volvamos a considerar uno de los ejemplos anteriores. Calcula la integral

$$\int_C \vec{F} \cdot d\vec{s}$$

si $\vec{F}(x, y) = e^y \vec{i} - \sin(\pi x) \vec{j}$ donde esta vez C representa el triángulo con vértices $(1, 0)$, $(0, 1)$, $(-1, 0)$ recorrido en el sentido de las agujas de un reloj.

13.3. LA INTEGRAL DE LÍNEA.

Solución:

Como ya habíamos visto, el triángulo está formado por tres segmentos de recta, que teniendo en cuenta la dirección habíamos parametrizado usando tres funciones $\vec{\sigma}_i$, con $i = 1, 2, 3$. Ahora, teniendo en cuenta la nueva orientación, podemos parametrizar los tres lados definiendo

$$C_1 : \vec{\sigma}_1(t) = t\vec{i} + (1-t)\vec{j}, \quad t \in [0, 1],$$

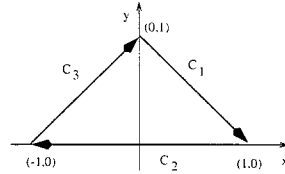
$$x_1(t) = t, \quad y_1(t) = 1-t;$$

$$C_2 : \vec{\sigma}_2(t) = (t-1)\vec{i} + t\vec{j}, \quad t \in [0, 1],$$

$$x_2(t) = t-1, \quad y_2(t) = t;$$

$$C_3 : \vec{\sigma}_3(t) = (1-2t)\vec{i}, \quad t \in [0, 1],$$

$$x_3(t) = 1-2t, \quad y_3(t) = 0.$$



Luego tendremos

$$\begin{aligned} \int_{C_1} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_1} \vec{F}(\vec{\sigma}_1(t)) \cdot \vec{\sigma}'_1(t) dt = \int_0^1 [e^{y_1(t)} x'_1(t) - \text{sen}(\pi x_1(t)) y'_1(t)] dt = \\ &= \int_0^1 [e^{(1-t)} + \text{sen}(\pi t)] dt = -\left(1 - e - \frac{2}{\pi}\right); \end{aligned}$$

$$\begin{aligned} \int_{C_2} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_2} \vec{F}(\vec{\sigma}_2(t)) \cdot \vec{\sigma}'_2(t) dt = \int_0^1 [e^{y_2(t)} x'_2(t) - \text{sen}(\pi x_2(t)) y'_2(t)] dt = \\ &= \int_0^1 [e^t - \text{sen}(\pi(t-1))] dt = -\left(1 - e - \frac{2}{\pi}\right); \end{aligned}$$

$$\begin{aligned} \int_{C_3} \vec{F} \cdot d\vec{s} &= \int_{\vec{\sigma}_3} \vec{F}(\vec{\sigma}_3(t)) \cdot \vec{\sigma}'_3(t) dt = \int_0^1 [e^{y_3(t)} x'_3(t) - \text{sen}(\pi x_3(t)) y'_3(t)] dt = \\ &= -\int_0^1 2 dt = -2. \end{aligned}$$

La integral sobre el triángulo es la suma de estas integrales:

$$\int_C \vec{F} \cdot d\vec{s} = -\left(1 - e - \frac{2}{\pi}\right) - \left(1 - e - \frac{2}{\pi}\right) - 2 = -4 + 2e + \frac{4}{\pi} \approx 2.71 = -\int_{C_{op}} \vec{F} \cdot d\vec{s}.$$

El teorema 13.3 nos dice que la integral de línea es una integral orientada, ya que ocurre un cambio de signo si se invierte la orientación de la curva. La integral de trayectoria no tiene esta propiedad. Esto se sigue del hecho de que al cambiar t por $-t$ (inversión de la orientación) sólo se cambia el signo de $\vec{\sigma}'(t)$, no su longitud. El siguiente teorema, que se puede demostrar de la misma manera que el anterior, afirma que las integrales de trayectoria no cambian bajo reparametrizaciones.

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Teorema 13.4 Sea $\bar{\sigma}$ una trayectoria \mathcal{C}^1 . $\bar{\sigma}: [a_1, b_1] \rightarrow \mathbb{R}^3$ y sea $\bar{\rho}: [a, b] \rightarrow \mathbb{R}^3$ cualquier reparametrización de $\bar{\sigma}$, invierta o no la orientación. Sea f una función con valores reales continua definida en la imagen de $\bar{\sigma}$. Entonces,

$$\int_{\bar{\sigma}} f(x, y, z) ds = \int_{\bar{\rho}} f(x, y, z) ds. \quad (13.14)$$

Consideraremos a continuación una técnica útil para evaluar integrales de línea cuando el campo vectorial dado \vec{F} es un **campo vectorial gradiente**, es decir, si se puede escribir $\vec{F} = \vec{\nabla} f$ para alguna función de valores reales f .

Teorema 13.5 Sea $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ una función de clase \mathcal{C}^1 , y sea $\bar{\sigma}: [a, b] \rightarrow \mathbb{R}^3$ una trayectoria \mathcal{C}^1 a trozos. Entonces,

$$\int_{\bar{\sigma}} \vec{\nabla} f(x, y, z) \cdot d\bar{s} = f(\bar{\sigma}(b)) - f(\bar{\sigma}(a)). \quad (13.15)$$

Demostración:

Aplicamos la regla de la cadena, teorema 10.6, a la función compuesta

$$F: t \rightsquigarrow f(\bar{\sigma}(t)) \quad \Rightarrow \quad F'(t) = (f \circ \bar{\sigma})'(t) = \vec{\nabla} f(\bar{\sigma}(t)) \cdot \bar{\sigma}'(t).$$

La función F es una función con valores reales de la variable t , de modo que por el teorema fundamental del cálculo, teorema 3.2 o regla de Barrow (3.1),

$$\int_a^b F'(t) dt = F(x) \Big|_a^b = F(b) - F(a) = f(\bar{\sigma}(b)) - f(\bar{\sigma}(a)).$$

Por lo tanto,

$$\int_{\bar{\sigma}} \vec{\nabla} f \cdot d\bar{s} = \int_a^b \vec{\nabla} f(\bar{\sigma}(t)) \cdot \bar{\sigma}'(t) dt = \int_a^b F'(t) dt = f(\bar{\sigma}(b)) - f(\bar{\sigma}(a)).$$

c.q.d

Lo que nos dice este teorema es que la integral sólo depende de los extremos de la curva y no de la propia curva C .

Obviamente, si podemos identificar el integrando como un gradiente, la evaluación de la integral será mucho más fácil. En el cálculo de una variable, toda integral es, en principio, obtenible hallando una primitiva. Sin embargo, para campos vectoriales esto no siempre es cierto, pues el campo vectorial no necesariamente es un gradiente.

Ejemplo 1. Integra el campo vectorial $\vec{F}(x, y) = (y^2, 2xy - e^y)$ sobre el arco de circunferencia $\bar{\sigma}(t) = (\cos t, \sin t)$, con $t \in [0, \pi/2]$.

13.3. LA INTEGRAL DE LÍNEA.

Solución:

Primero intentemos comprobar si \vec{F} es un gradiente. Observamos que \vec{F} tiene la forma $P(x, y)\vec{i} + Q(x, y)\vec{j}$ con

$$P(x, y) = y^2 \quad \text{y} \quad Q(x, y) = 2xy - e^y.$$

Dado que P y Q son continuamente diferenciables en todo punto y que

$$\frac{\partial P}{\partial y}(x, y) = 2y = \frac{\partial Q}{\partial x}(x, y)$$

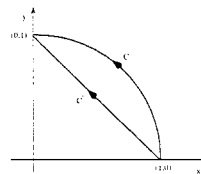
usando lo visto en la sección 13.1, podemos concluir que \vec{F} es un gradiente.

Dado que la integral sólo depende de los extremos de la curva y no de la propia curva C , podemos simplificar el cálculo integrando a lo largo de cualquier segmento C' que une los extremos.

Parametrizamos C' haciendo

$$C' : \quad \vec{\sigma}(t) = (1-t)\vec{i} + t\vec{j}, \quad t \in [0, 1],$$

$$x(t) = 1-t, \quad y(t) = t.$$



Tenemos entonces

$$\begin{aligned} \int_C \vec{F} \cdot d\vec{s} &= \int_{C'} \vec{F} \cdot d\vec{s} = \int_0^1 \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt = \\ &= \int_0^1 [y^2(t)x'(t) + [2x(t)y(t) - e^{y(t)}]y'(t)] dt = \\ &= \int_0^1 [2t - 3t^2 - e^t] dt = 1 - e. \end{aligned}$$

Una vez comprobado que \vec{F} es un campo gradiente, hay una solución alternativa a nuestro problema. Determinemos la función escalar $f(x, y)$ con los métodos de la sección 13.1. Dado que

$$P(x, y) = \frac{\partial f}{\partial x} = y^2 \quad \text{y} \quad Q(x, y) = \frac{\partial f}{\partial y} = 2xy - e^y,$$

tenemos

$$f(x, y) = xy^2 + \phi(y) \quad \text{luego} \quad \frac{\partial f}{\partial y} = 2xy + \phi'(y).$$

Las dos expresiones para $\frac{\partial f}{\partial y}$ sólo coinciden si

$$\phi'(y) = -e^y \quad \text{luego si} \quad \phi(y) = -e^y + c.$$

Esto significa que $f(x, y) = xy^2 - e^y - c$. Dado que la curva C comienza en $(1, 0)$ y acaba en $(0, 1)$, por el teorema 13.5 y usando la fórmula (13.15), vemos que

$$\int_C \vec{F} \cdot d\vec{s} = f(0, 1) - f(1, 0) = (-e + c) - (-1 + c) = 1 - e.$$

Ejemplo 2. Sea $\vec{\sigma}$ la trayectoria dada por

$$\vec{\sigma}(t) = \left(\frac{t^4}{4}, \sin^3 \left(\frac{\pi t}{2} \right), 0 \right), \quad \text{con } t \in [0, 1]$$

Calcula la integral

$$\int_{\vec{\sigma}} y \, dx + x \, dy.$$

Solución:

Aquí tenemos la integral de una forma diferencial, por lo que podemos reconocer el campo vectorial $\vec{F}(x, y, z) = (y, x, 0)$ como el gradiente de la función $f(x, y, z) = xy$. Así, usando la fórmula (13.15)

$$\int_{\vec{\sigma}} y \, dx + x \, dy = f(\vec{\sigma}(1)) - f(\vec{\sigma}(0)) = \frac{1}{4} \cdot \sin^3 \left(\frac{\pi}{2} \right) - \frac{0}{4} \cdot \sin^3 \left(\frac{0}{2} \right) = \frac{1}{4}.$$

Los campos gradientes son importantes en problemas físicos. Usualmente, $V = -f$ representa un potencial de energía (gravitacional, eléctrico, etc.) y \vec{F} representa una fuerza (si se usa el signo menos, entonces V es decreciente en dirección de \vec{F} ; así, una partícula sobre la que actúe \vec{F} se mueve en la dirección en la que decrezca el potencial). Consideremos el ejemplo de una partícula de masa m en el campo gravitatorio de la Tierra; en este caso se toma f como $\frac{GmM}{r}$ o $\vec{F} = -\frac{GmM}{r} \vec{r}$, donde G es la constante gravitacional, M es la masa de la Tierra y r es la distancia al centro de la Tierra. La fuerza correspondiente es $\vec{F} = \frac{GmM}{r^3} \vec{r} = \frac{GmM}{r^2} \vec{n}$, donde \vec{n} es el vector radial unitario (estudiaremos este caso más adelante.) Nótese que \vec{F} no está definido en el punto $r = 0$.

Deseamos caracterizar los campos vectoriales que se pueden escribir como un gradiente. Nuestra labor se simplifica de manera considerable gracias al teorema de Stokes, teoremas 13.14 y 13.15, que estudiaremos más adelante y que ahora damos por válidos.

Teorema 13.6 *Sea \vec{F} un campo vectorial C^1 definido en \mathbb{R}^3 excepto, quizás, en un número finito de puntos. Las siguientes propiedades sobre \vec{F} son equivalentes:*

(i) *Para cualquier curva cerrada simple orientada C ,*

$$\oint_C \vec{F} \cdot d\vec{s} = 0.$$

(ii) *Para cualesquiera dos curvas simples orientadas C_1 y C_2 que tengan los mismos extremos*

$$\int_{C_1} \vec{F} \cdot d\vec{s} = \int_{C_2} \vec{F} \cdot d\vec{s}.$$

(iii) *\vec{F} es el gradiente de alguna función f ; esto es, $\vec{F} = \vec{\nabla} f$ (y si \vec{F} tiene un punto excepcional donde no está definido, tampoco f está definido ahí).*

13.3. LA INTEGRAL DE LÍNEA.

$$(iv) \quad \vec{\nabla} \times \vec{F} = \vec{0}.$$

Un campo vectorial que satisfaga una (y por lo tanto, todas) de las condiciones (i) – (iv) se llama **campo vectorial conservativo**.

Demostración:

Probaremos la siguiente cadena de implicaciones, lo cual probará el teorema:

$$(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i).$$

Primero demostraremos que la condición (i) implica la condición (ii). Supongamos que $\vec{\sigma}_1$ y $\vec{\sigma}_2$ sean parametrizaciones que representan a C_1 y C_2 , con los mismos extremos. Construimos la curva cerrada $\vec{\sigma}$, obtenida recorriendo primero $\vec{\sigma}_1$ y después $\vec{\sigma}_2$, o, simbólicamente, la curva $\vec{\sigma} = \vec{\sigma}_1 - \vec{\sigma}_2$. Suponiendo que $\vec{\sigma}$ es simple, la condición (i) da

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_{\vec{\sigma}_1} \vec{F} \cdot d\vec{s} - \int_{\vec{\sigma}_2} \vec{F} \cdot d\vec{s} = 0.$$

de modo que se cumple la condición (ii).

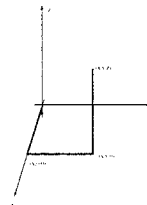
A continuación probaremos que la condición (ii) implica la condición (iii). Sea C cualquier curva orientada simple, que une a un punto como $(0, 0, 0)$ con (x, y, z) y supongamos que C esté representada por la parametrización $\vec{\sigma}$ (si $(0, 0, 0)$ es el punto excepcional de \vec{F} , podemos escoger un punto de inicio de $\vec{\sigma}$ diferente sin que quede afectada la argumentación).

Definimos $f(x, y, z)$ como la integral $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}$.

Por la hipótesis (ii), $f(x, y, z)$ es independiente de la curva C .

Mostraremos que $\vec{F} = \text{grad } f$.

En efecto, escojamos $\vec{\sigma}$ como la trayectoria mostrada en la figura, de modo que



$$f(x, y, z) = \int_0^x F_1(t, 0, 0) dt + \int_0^y F_2(x, t, 0) dt + \int_0^z F_3(x, y, t) dt$$

donde $\vec{F} = (F_1, F_2, F_3)$. Se sigue de manera inmediata, que $\partial f / \partial z = F_3$. Permutando x , y y z , podemos demostrar similarmente que $\partial f / \partial x = F_1$ y $\partial f / \partial y = F_2$, esto es, $\vec{\nabla} f = \vec{F}$. Tercero, la condición (iii) implica la condición (iv), pues ya se dijo en la identidad 14 dada en la sección 13.1.2 que

$$\vec{\nabla} \times \vec{\nabla} f = \vec{0}.$$

Finalmente, sean $\vec{\sigma}$ una representación de una curva cerrada C y S cualquier superficie cuya frontera sea $\vec{\sigma}$ (si \vec{F} tiene puntos excepcionales, se escoge S de tal manera que los evite). Por el teorema de Stokes, teoremas 13.14 y 13.15,

$$\int_C \vec{F} \cdot d\vec{s} = \int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_S (\vec{\nabla} \times \vec{F}) \cdot \vec{n} dS.$$

Como $\vec{\nabla} \times \vec{F} = \vec{0}$, esta integral se anula, de modo que la condición (iv) \Rightarrow condición (i).
c.q.d.

Hay varias interpretaciones físicas útiles de $\int_C \vec{F} \cdot d\vec{s}$. Ya vimos que una es el trabajo realizado por \vec{F} al mover una partícula a lo largo de C . Otra interpretación es el concepto de “circulación”, que veremos al final de la sección 13.8. En este caso pensamos en \vec{F} como el campo de velocidad de un fluido; esto es, a cada punto P en el espacio, \vec{F} asigna el vector velocidad del fluido en P . Si tomamos C como una curva cerrada, y Δs es una pequeña cuerda dirigida de C , entonces $\vec{F} \cdot \Delta s$ es aproximadamente la componente tangencial de \vec{F} por $\|\Delta s\|$. La integral $\int_C \vec{F} \cdot d\vec{s}$ es la componente neta tangencial alrededor de C . Esto significa que si colocamos una pequeña rueda con aspas en el fluido, girará si la circulación del fluido fuera diferente de cero, o $\int_C \vec{F} \cdot d\vec{s} \neq 0$. Así, con frecuencia nos referimos a la integral de línea $\int_C \vec{F} \cdot d\vec{s}$ como la **circulación de \vec{F} alrededor de C** . Por el teorema 13.6, un campo \vec{F} no tiene circulación si y sólo si $\text{rot } \vec{F} = \vec{0}$. De aquí, que un campo vectorial F con $\text{rot } \vec{F} = \vec{0}$ se llame **irrotacional**. Hemos probado entonces que un campo vectorial en \mathbb{R}^3 es irrotacional si y sólo si es el campo gradiente de alguna función: esto es, si y sólo si $\vec{F} = \vec{\nabla} f$. La función f se llama **potencial para \vec{F}** .

Ejemplo. Sea \vec{F} el campo vectorial en \mathbb{R}^3 definido por

$$\vec{F}(x, y, z) = (y, z \cos(yz) + x, y \cos(yz)).$$

Demuestra que \vec{F} es irrotacional y halla un potencial escalar para \vec{F} .

Solución:

Calculamos $\vec{\nabla} \times \vec{F}$:

$$\text{rot } \vec{F} = \vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y & z \cos(yz) + x & y \cos(yz) \end{vmatrix} = \vec{0}.$$

de modo que \vec{F} es irrotacional. Podemos hallar un potencial escalar de varias maneras. Método 1. Con la técnica usada en el teorema 13.6 para probar que la condición (ii) implica la condición (iii), podemos hacer

$$\begin{aligned} f(x, y, z) &= \int_0^x F_1(t, 0, 0) dt + \int_0^y F_2(x, t, 0) dt + \int_0^z F_3(x, y, t) dt = \\ &= \int_0^x 0 dt + \int_0^y x dt + \int_0^z y \cos(yt) dt = xy + \text{sen}(yz). \end{aligned}$$

Se verifica fácilmente que, como se requiere, $\vec{\nabla} f = \vec{F}$:

$$\vec{\nabla} f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} + \frac{\partial f}{\partial z} \vec{k} = y \vec{i} + (x + z \cos(yz)) \vec{j} + y \cos(yz) \vec{k} = \vec{F}.$$

Método 2. Como sabemos que existe f , sabemos que es posible resolver

$$\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x + z \cos(yz), \quad \frac{\partial f}{\partial z} = y \cos(yz).$$

para $f(x, y, z)$. Éstas son equivalentes a las ecuaciones simultáneas

13.3. LA INTEGRAL DE LÍNEA.

- (a) $f(x, y, z) = xy + h_1(y, z)$,
 (b) $f(x, y, z) = \text{sen}(yz) + xy + h_2(x, z)$,
 (c) $f(x, y, z) = \text{sen}(yz) + h_3(x, y)$,

para funciones h_1 , h_2 y h_3 , independientes de x , y y z (respectivamente). Para deducir de manera más sistemática la fórmula notamos que como

$$f(x, y, z) = xy + h_1(y, z) \quad \text{y} \quad \frac{\partial f}{\partial z} = y \cos(yz),$$

tenemos

$$\frac{\partial h_1(y, z)}{\partial z} = y \cos(yz) \quad \Rightarrow \quad h_1(y, z) = \int y \cos(yz) dz + g(y) = \text{sen}(yz) + g(y).$$

Por lo tanto, sustituyendo esto en la ecuación (a) obtenemos

$$f(x, y, z) = xy + \text{sen}(yz) + g(y),$$

pero por la ecuación (b),

$$g(y) = h_2(x, z).$$

Debido a que el lado derecho de esta ecuación es una función de x y z y el lado izquierdo es una función sólo de y , concluimos que deben ser iguales a alguna constante C . Así,

$$f(x, y, z) = xy + \text{sen}(yz) + C$$

y hemos determinado f salvo por una constante.

Si $\vec{F} = P\vec{i} + Q\vec{j}$, entonces $\vec{\nabla} \times \vec{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \vec{k}$, de modo que la condición $\vec{\nabla} \times \vec{F} = \vec{0}$ se reduce a $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$. Así, tenemos el siguiente corolario

Teorema 13.7 Si \vec{F} es un campo vectorial C^1 en \mathbb{R}^2 de la forma $\vec{F} = P\vec{i} + Q\vec{j}$ con $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$, entonces $\vec{F} = \vec{\nabla} f$ para alguna f definida en \mathbb{R}^2 .

Este corolario puede ser falso si \vec{F} deja de ser de clase C^1 incluso en un solo punto. Sin embargo, en \mathbb{R}^3 se permiten excepciones en varios puntos (ver el teorema 13.6).

Ejemplo. Determina si los campos vectoriales

- a) $\vec{F} = e^{xy}\vec{i} + e^{x+y}\vec{j}$, b) $\vec{F} = (2x \cos y)\vec{i} - (x^2 \text{sen } y)\vec{j}$,

son campos gradientes.

Solución:

(a) En este caso $P(x, y) = e^{xy}$ y $Q(x, y) = e^{x-y}$, calculando

$$\frac{\partial P}{\partial y} = x e^{xy}, \quad \frac{\partial Q}{\partial x} = e^{x-y},$$

vemos que no son iguales, de modo que \vec{F} no puede tener una función potencial.

(b) En este caso $P(x, y) = 2x \cos y$ y $Q(x, y) = -x^2 \operatorname{sen} y$ hallando

$$\frac{\partial P}{\partial y} = -2x \operatorname{sen} y = \frac{\partial Q}{\partial x}.$$

vemos que \vec{F} tiene una función potencial f . Para calcular f resolvemos las ecuaciones

$$\frac{\partial f}{\partial x} = 2x \cos y, \quad \frac{\partial f}{\partial y} = -x^2 \operatorname{sen} y.$$

Así,

$$f(x, y) = x^2 \cos y + h_1(y), \quad f(x, y) = x^2 \cos y + h_2(x).$$

Si h_1 y h_2 son la misma constante, entonces se satisfacen ambas ecuaciones, de modo que $f(x, y) = x^2 \cos y + C$ es un potencial para \vec{F} .

Concluimos esta sección con un teorema que es bastante parecido en esencia al teorema 13.6. El teorema 13.6 fue motivado, en parte como un recíproco al resultado $\operatorname{rot} \nabla f = \vec{0}$ para cualquier función C^1 , $f: \mathbb{R}^3 \rightarrow \mathbb{R}$, esto es, si $\operatorname{rot} \vec{F} = \vec{0}$, entonces $\vec{F} = \nabla f$. También sabemos que $\operatorname{div}(\operatorname{rot} \vec{G}) = 0$ para cualquier campo vectorial \vec{G} de clase C^2 . Podríamos plantearnos la validez del enunciado recíproco: si $\operatorname{div} \vec{F} = 0$ ¿es \vec{F} el rotacional de un campo vectorial \vec{G} ? La respuesta, afirmativa, la da el siguiente teorema.

Teorema 13.8 *Si \vec{F} es un campo vectorial C^1 en \mathbb{R}^3 con $\operatorname{div} \vec{F} = 0$, entonces existe un campo vectorial \vec{G} de clase C^1 tal que $\vec{F} = \operatorname{rot} \vec{G}$.*

Nótese que a diferencia de \vec{F} del teorema 13.6, al campo vectorial \vec{F} de este último teorema no se le permite tener un punto excepcional.

Concluimos esta sección con

Algunas aplicaciones de las integrales de línea.

Si C es una curva de \mathbb{R}^3 que representa un alambre muy fino de densidad variable $f(x, y, z)$ en el punto $(x, y, z) \in C$ entonces:

a) La **masa** del alambre es:

$$m(C) = \int_C f \, ds. \tag{13.16}$$

b) Las **coordenadas del centro de gravedad** del alambre, $(\bar{x}, \bar{y}, \bar{z})$, son:

$$\bar{x} = \frac{\int_C x f(x, y, z) \, ds}{m(C)}, \quad \bar{y} = \frac{\int_C y f(x, y, z) \, ds}{m(C)}, \quad \bar{z} = \frac{\int_C z f(x, y, z) \, ds}{m(C)}. \tag{13.17}$$

13.3. LA INTEGRAL DE LÍNEA.

c) Si L es una recta en \mathbb{R}^3 y $\delta(x, y, z)$ es la distancia del punto $(x, y, z) \in C$ a la recta L , el **momento de inercia** del alambre respecto a L es:

$$I_L = \int_C \delta(x, y, z)^2 f(x, y, z) ds.$$

En particular los momentos de inercia respecto a los ejes coordenados son:

$$\begin{aligned} I_x &= \int_C (y^2 + z^2) f(x, y, z) ds \\ I_y &= \int_C (x^2 + z^2) f(x, y, z) ds \\ I_z &= \int_C (x^2 + y^2) f(x, y, z) ds. \end{aligned} \tag{13.18}$$

Ejercicios.

- 1) Sea $\vec{\sigma}$ una trayectoria suave. Supongamos que \vec{F} es ortogonal a $\vec{\sigma}'(t)$ en $\vec{\sigma}(t)$. Demuestra que

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = 0.$$

- 2) Sea $\vec{F} = \vec{\nabla} f$ para una función escalar dada. Sea $\vec{\sigma}(t)$ una curva cerrada, esto es $\vec{\sigma}(a) = \vec{\sigma}(b)$. Demuestra que

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = 0.$$

- 3) Integra los campos vectoriales $\vec{F}(x, y) = y\vec{i} + x\vec{j}$ y $\vec{G}(x, y) = x\vec{i} + y\vec{j}$ sobre las curvas indicadas:

- a) $\vec{\sigma}(t) = t\vec{i} + t^2\vec{j}$, con $t \in [0, 1]$;
 b) $\vec{\sigma}(t) = t^3\vec{i} - 2t\vec{j}$, con $t \in [0, 1]$.

- 4) Integra los campos vectoriales $\vec{F}(x, y, z) = yz\vec{i} + x^2\vec{j} + xz\vec{k}$ y $\vec{G}(x, y, z) = e^x\vec{i} + e^y\vec{j} + e^z\vec{k}$ sobre las curvas indicadas:

- a) el segmento que une el punto $(0, 0, 0)$ a $(1, 1, 1)$;
 b) $\vec{\sigma}(t) = t\vec{i} + t^2\vec{j} + t^3\vec{k}$, con $t \in [0, 1]$.

- 5) Calcula el trabajo realizado por la fuerza $\vec{F}(x, y, z) = x\vec{i} + xy\vec{j} + xyz\vec{k}$ aplicada a un punto material que se mueve a lo largo de una línea recta desde $(0, 1, 4)$ hasta $(1, 0, -4)$.

- 6) Una masa puntual m se mueve desde el instante $t = 0$ hasta $t = 1$ de tal forma que su posición venga dada por la función vectorial $\vec{\sigma}(t) = \alpha t\vec{i} + \beta t^2\vec{j} + \gamma t^3\vec{k}$. Halla la fuerza total que actúa sobre dicha masa en el instante t y calcula el trabajo realizado por esta fuerza en el intervalo de tiempo $[0, 1]$.

7) Evalúa la integral

$$\int_{\vec{\sigma}} y \, dy + (3y^3 - x) \, dy + z \, dz$$

para cada una de las trayectorias $\vec{\sigma}(t) = (t, t^n, 0)$, para $0 \leq t \leq 1$ y $n = 1, 2, 3, \dots$

8) Evalúa la integral de línea del campo vectorial $\vec{F}(x, y) = (x, y)$ alrededor de la hipocicloide $\vec{\sigma}(t) = (\cos^3 t, \sin^3 t)$, para $0 \leq t \leq 2\pi$.

9) Sea \vec{a} un vector constante y $\vec{F} = \vec{a} \times \vec{r}$ [como siempre $\vec{r}(x, y, z) = (x, y, z)$]. ¿Es \vec{F} conservativo? De ser así, halla su función potencial.

13.4 Superficies parametrizadas.

En las secciones 13.2 y 13.3 hemos estudiado integrales de funciones escalares y vectoriales a lo largo de curvas. Pasemos ahora a integrales sobre superficies y comencemos por estudiar la geometría de las superficies.

Ya conocemos un tipo de superficie, esto es, la gráfica de una función $f(x, y)$ del tipo estudiado en el capítulo 10 para las cuales sabemos también calcular los planos tangentes. Sin embargo, no podemos limitarnos sólo a este tipo de superficies. Muchas se representan como superficies de nivel de funciones. Por ejemplo, una hoja que se dobla sobre sí misma no es la gráfica de una función porque para cada $(x_0, y_0) \in \mathbb{R}^2$ debe haber un único z_0 con (x_0, y_0, z_0) en la gráfica. De la misma manera un toroide no es la gráfica de una función de la forma $z = f(x, y)$.

Estas observaciones nos impulsan a extender nuestra definición de superficie, y como en el caso de las curvas, queremos hacer una distinción entre una función, que es la parametrización, y su imagen, que es el objeto geométrico.

Definición. Una **superficie parametrizada** es una función $\vec{\Phi}: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$, donde D es algún dominio en \mathbb{R}^2 . La **superficie** S correspondiente a la función $\vec{\Phi}$ es su imagen $S = \vec{\Phi}(D)$. Podemos escribir

$$\vec{\Phi}(u, v) = (x(u, v), y(u, v), z(u, v)).$$

Si $\vec{\Phi}$ es diferenciable o es de clase \mathcal{C}^1 , (las funciones componentes $x(u, v)$, $y(u, v)$, $z(u, v)$ son funciones diferenciables o de clase \mathcal{C}^1), llamamos a S **superficie diferenciable** o \mathcal{C}^1 .

Ejemplos.

- La gráfica de una función $y = f(x)$, con $x \in [a, b]$ puede parametrizarse haciendo $\vec{r}(u) = u\vec{i} + f(u)\vec{j}$. De manera parecida, la gráfica de una función $z = f(x, y)$, con $(x, y) \in \Omega \subset \mathbb{R}^2$ puede parametrizarse haciendo

$$\vec{r}(u, v) = u\vec{i} + v\vec{j} + f(u, v)\vec{k}.$$

13.4. SUPERFICIES PARAMETRIZADAS.

- Si dos vectores \vec{a} y \vec{b} no son paralelos, el conjunto de todas las combinaciones lineales $u\vec{a} + v\vec{b}$ genera un plano que pasa por el origen. Podemos parametrizar este plano haciendo $\vec{r}(u, v) = u\vec{a} + v\vec{b}$, con $(u, v) \in \mathbb{R}^2$. El nuevo plano paralelo al anterior y que pasa por el extremo de \vec{c} puede parametrizarse haciendo

$$\vec{r}(u, v) = u\vec{a} + v\vec{b} + \vec{c}, \quad \text{con } (u, v) \in \mathbb{R}^2.$$

Este plano claramente contiene las dos rectas $\vec{r}(u, 0) = u\vec{a} + \vec{c}$ y $\vec{r}(0, v) = v\vec{b} + \vec{c}$.

- La esfera de radio a centrada en el origen puede parametrizarse, usando las coordenadas esféricas, como

$$\vec{r}(\theta, \phi) = a \cos \theta \sin \phi \vec{i} + a \sin \theta \sin \phi \vec{j} + a \cos \phi \vec{k}, \quad 0 \leq \theta < 2\pi, 0 \leq \phi \leq \pi.$$

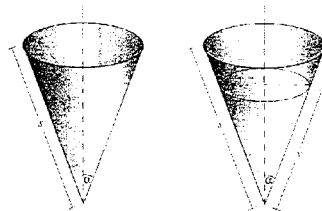
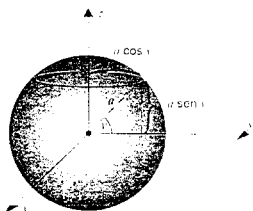
Pero también puede ser parametrizada haciendo

$$\vec{r}(u, v) = a \cos u \cos v \vec{i} + a \sin u \cos v \vec{j} + a \sin v \vec{k}, \quad 0 \leq u \leq 2\pi, -\frac{\pi}{2} \leq v \leq \frac{\pi}{2}.$$

Para demostrar esta parametrización, notamos que los puntos de latitud v forman una circunferencia de radio $a \cos v$ en el plano horizontal $z = a \sin v$. Esta circunferencia puede parametrizarse poniendo

$$\vec{R}(u) = a \cos v (\cos u \vec{i} + \sin u \vec{j}) + a \sin v \vec{k}, \quad \text{con } u \in [0, 2\pi].$$

Desarrollando esta expresión y haciendo variar v desde $-\pi/2$ hasta $\pi/2$ obtenemos la esfera entera. Es fácil comprobar que la parametrización satisface la ecuación cartesiana de la esfera $x^2 + y^2 + z^2 = a^2$.



- Consideremos ahora un cono con semiángulo en el vértice igual a α y longitud de generatriz igual a s . Los puntos de la generatriz a distancia v del vértice forman una circunferencia de radio $v \sin \alpha$ en el plano horizontal $z = v \cos \alpha$. Esta circunferencia puede parametrizarse haciendo

$$\vec{R}(u) = v \sin \alpha (\cos u \vec{i} + \sin u \vec{j}) + v \cos \alpha \vec{k}, \quad \text{con } u \in [0, 2\pi].$$

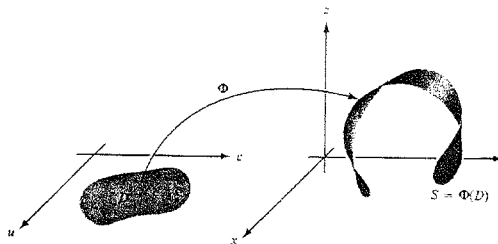
Dado que podemos obtener el cono entero haciendo variar v desde cero hasta s , el cono está parametrizado por

$$\vec{r}(u, v) = v \sin \alpha \cos u \vec{i} + v \sin \alpha \sin u \vec{j} + v \cos \alpha \vec{k}, \quad 0 \leq u \leq 2\pi, 0 \leq v \leq s.$$

Es fácil comprobar que la parametrización satisface la ecuación cartesiana del cono $x^2 + y^2 = z^2$.

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Volvamos a la superficie parametrizada por la aplicación $\vec{\Phi}(u, v)$. Podemos pensar que dicha aplicación $\vec{\Phi}$ tuerce o dobla la región D en el plano para producir la superficie S en el espacio.



Supongamos que $\vec{\Phi}$ es diferenciable en $(u_0, v_0) \in \mathbb{R}^2$. Si fijamos la variable u en u_0 obtenemos una función $\mathbb{R} \rightarrow \mathbb{R}^3$ dada por $t \rightsquigarrow \vec{\Phi}(u_0, t)$, cuya imagen es una curva sobre la superficie S .

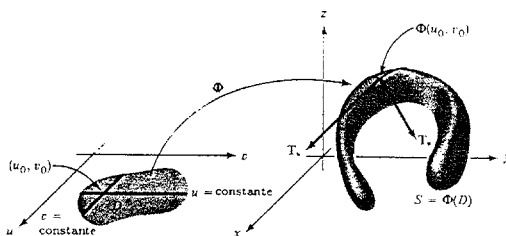
Por lo ya visto en temas anteriores, ver capítulo 10, el vector tangente a esta curva en el punto $\vec{\Phi}(u_0, v_0)$ está dado por

$$\vec{T}_v = \frac{\partial x}{\partial v}(u_0, v_0) \vec{i} + \frac{\partial y}{\partial v}(u_0, v_0) \vec{j} + \frac{\partial z}{\partial v}(u_0, v_0) \vec{k}.$$

De manera análoga, si fijamos la variable v en v_0 obtenemos una función $\mathbb{R} \rightarrow \mathbb{R}^3$ dada por $t \rightsquigarrow \vec{\Phi}(t, v_0)$, cuya imagen es otra curva sobre la superficie S y el vector tangente a esta curva en el punto $\vec{\Phi}(u_0, v_0)$ está dado por

$$\vec{T}_u = \frac{\partial x}{\partial u}(u_0, v_0) \vec{i} + \frac{\partial y}{\partial u}(u_0, v_0) \vec{j} + \frac{\partial z}{\partial u}(u_0, v_0) \vec{k}.$$

Como los vectores \vec{T}_v y \vec{T}_u son tangentes a dos curvas sobre la superficie en $\vec{\Phi}(u_0, v_0)$, deben determinar el plano tangente a la superficie en este punto; esto es, $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v$ debe ser normal a la superficie. Al producto vectorial $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v$ se le llama **producto vectorial fundamental** de la superficie.



Definición. Decimos que la superficie S es **suave** en $\vec{\Phi}(u_0, v_0)$ si $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v \neq \vec{0}$ en (u_0, v_0) ; esto es, existirá el plano tangente. La superficie es suave si es suave en todos los puntos $\vec{\Phi}(u_0, v_0) \in S$. Intuitivamente, una superficie suave no tiene “esquinas”.

13.4. SUPERFICIES PARAMETRIZADAS.

Ejemplo. Considera la superficie dada por las ecuaciones

$$x = u \cos v, \quad y = u \sin v, \quad z = u, \quad v > 0.$$

¿Es diferenciable esta superficie? ¿Es suave?

Solución:

Estas ecuaciones describen la superficie $z = \sqrt{x^2 + y^2}$ que es un cono. Esta superficie es un cono con “punta” en $(0, 0, 0)$; es una superficie diferenciable, pues cada función componente es diferenciable como función de u y v . Sin embargo, la superficie no es suave en $(0, 0, 0)$, siendo $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = \vec{0}$:

$$\vec{T}_u = \frac{\partial x}{\partial u}(0, 0) \vec{i} + \frac{\partial y}{\partial u}(0, 0) \vec{j} + \frac{\partial z}{\partial u}(0, 0) \vec{k} = \cos 0 \cdot \vec{i} + \sin 0 \cdot \vec{j} + 1 \cdot \vec{k} = \vec{i} + \vec{k},$$

$$\vec{T}_v = \frac{\partial x}{\partial v}(0, 0) \vec{i} + \frac{\partial y}{\partial v}(0, 0) \vec{j} + \frac{\partial z}{\partial v}(0, 0) \vec{k} = 0 \cdot (-\sin 0) \cdot \vec{i} + 0 \cdot (\cos 0) \cdot \vec{j} + 0 \cdot \vec{k} = \vec{0}.$$

Definición. Si una superficie parametrizada $\vec{\Phi}: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es **suave** en $\vec{\Phi}(u_0, v_0)$, esto es, si $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v \neq \vec{0}$ en (u_0, v_0) , definimos el **plano tangente** de la superficie en $\vec{\Phi}(u_0, v_0)$ como el plano determinado por los vectores \vec{T}_u y \vec{T}_v . Siendo $\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v$ un vector normal a la superficie, una ecuación del plano tangente en (x_0, y_0, z_0) a la superficie está dado por

$$(x - x_0, y - y_0, z - z_0) \cdot \vec{N}(u_0, v_0) = 0 \quad (13.19)$$

o de forma equivalente

$$n_1(x - x_0) + n_2(y - y_0) + n_3(z - z_0) = 0$$

donde (n_1, n_2, n_3) son las componentes del vector producto vectorial fundamental.

Ejemplo 1. Considera la superficie dada por las ecuaciones

$$x = u \cos v, \quad y = u \sin v, \quad z = u^2 + v^2.$$

¿Dónde existe un plano tangente? Halla el plano tangente en $\vec{\Phi}(1, 0)$.

Solución:

Calculamos

$$\vec{T}_u = \frac{\partial x}{\partial u}(u, v) \vec{i} + \frac{\partial y}{\partial u}(u, v) \vec{j} + \frac{\partial z}{\partial u}(u, v) \vec{k} = \cos v \vec{i} + \sin v \vec{j} + 2u \vec{k},$$

$$\vec{T}_v = \frac{\partial x}{\partial v}(u, v) \vec{i} + \frac{\partial y}{\partial v}(u, v) \vec{j} + \frac{\partial z}{\partial v}(u, v) \vec{k} = -u \sin v \vec{i} + u \cos v \vec{j} + 2v \vec{k}.$$

El plano tangente en $\vec{\Phi}(u, v)$ será el conjunto de vectores que pasan por $\vec{\Phi}(u, v)$ y son perpendiculares a

$$\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = (-2u^2 \cos v + 2v \sin v, -2u^2 \sin v - 2v \cos v, u),$$

si resulta ser no nulo. Claramente $\vec{N}(u, v) = \vec{0}$ en $(u, v) = (0, 0)$, esto es no existe plano tangente en $\vec{\Phi}(0, 0) = (0, 0, 0)$. Sin embargo, en todo punto $(u, v) \neq (0, 0)$ tenemos $\vec{N}(u, v) \neq \vec{0}$ y entonces existe el plano tangente. En particular, en $\vec{\Phi}(1, 0) = (1, 0, 1)$, usando la ecuación (13.19), tenemos que

$$\vec{N}(1, 0) = (-2, 0, 1) \Rightarrow (x - 1, y - 0, z - 1) \cdot (-2, 0, 1) = 0 \Rightarrow z = 2x - 1,$$

es la ecuación del plano tangente.

Ejemplo 2. Sea S una superficie obtenida como gráfica de función diferenciable $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. Demuestra que la superficie es suave en todos los puntos $(u_0, v_0, f(u_0, v_0)) \in \mathbb{R}^3$.

Solución:

Escribamos S en forma paramétrica como

$$x = u, \quad y = v, \quad z = f(u, v), \quad \text{que es lo mismo que } z = f(x, y).$$

Entonces

$$\begin{aligned} \vec{T}_u &= \frac{\partial x}{\partial u}(u, v) \vec{i} + \frac{\partial y}{\partial u}(u, v) \vec{j} + \frac{\partial z}{\partial u}(u, v) \vec{k} = \vec{i} + \frac{\partial f}{\partial u}(u_0, v_0) \vec{k}, \\ \vec{T}_v &= \frac{\partial x}{\partial v}(u, v) \vec{i} + \frac{\partial y}{\partial v}(u, v) \vec{j} + \frac{\partial z}{\partial v}(u, v) \vec{k} = \vec{j} + \frac{\partial f}{\partial v}(u_0, v_0) \vec{k}. \end{aligned}$$

Para todo $(u_0, v_0) \in \mathbb{R}^2$ tenemos

$$\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = \left(-\frac{\partial f}{\partial u}(u_0, v_0), -\frac{\partial f}{\partial v}(u_0, v_0), 1 \right) \neq \vec{0}.$$

Luego la parametrización es suave en todos los puntos.

El plano tangente en $\vec{\Phi}(u_0, v_0) = (x_0, y_0, z_0) = (u_0, v_0, f(u_0, v_0))$ será el conjunto de vectores que pasan por $\vec{\Phi}(u_0, v_0)$ y perpendiculares a $\vec{N}(u_0, v_0) = \vec{T}_u \times \vec{T}_v$, esto es, usando la ecuación (13.19),

$$(x - x_0, y - y_0, z - z_0) \cdot \left(-\frac{\partial f}{\partial u}(u_0, v_0), -\frac{\partial f}{\partial v}(u_0, v_0), 1 \right) = 0$$

o lo que es lo mismo

$$z - z_0 = \frac{\partial f}{\partial u}(u_0, v_0)(x - x_0) + \frac{\partial f}{\partial v}(u_0, v_0)(y - y_0)$$

es la ecuación del plano tangente. Esta última ecuación coincide con la definición de plano tangente (10.1) obtenida cuando la superficie S es la gráfica de una función.

Es oportuno considerar el problema de calcular el área de una superficie, antes de pasar a estudiar la integración de funciones sobre superficies. De aquí en adelante consideraremos sólo superficies suaves a trozos que sean uniones de imágenes de superficies parametrizadas $\vec{\Phi}_i: D_i \rightarrow \mathbb{R}^3$ para las cuales D_i es una región elemental del plano, $\vec{\Phi}_i$ es de clase \mathcal{C}^1 y uno

13.4. SUPERFICIES PARAMETRIZADAS.

a uno, excepto, quizá, en la frontera de D_i y S_i , la imagen de $\vec{\Phi}_i$ es suave, excepto, quizá, en un número finito de puntos.

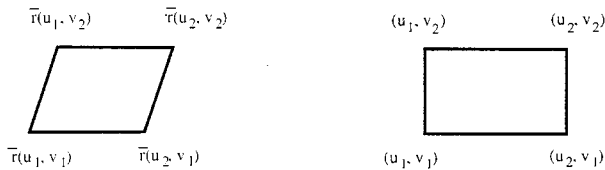
Antes de dar la definición de área de superficie de una superficie parametrizada y la justificación en términos de sumas de Riemann, consideremos el caso más simple del área de un paralelogramo.

Ya hemos visto que una función lineal del tipo $\vec{r}(u, v) = u\vec{a} + v\vec{b} + \vec{c}$, con \vec{a} y \vec{b} no paralelos, parametriza un plano p . Las rectas horizontales del plano uv , con ecuación $v = v_0$, se transforman en rectas paralelas al vector \vec{a} , esto es $\vec{r}(u, v_0) = u\vec{a} + v_0\vec{b} + \vec{c}$ tiene por vector director el vector \vec{a} ; y las rectas verticales del plano uv , con ecuación $u = u_0$, se transforman en rectas paralelas al vector \vec{b} , esto es $\vec{r}(u_0, v) = v\vec{b} + u_0\vec{a} + \vec{c}$ tiene por vector director el vector \vec{b} . Luego un rectángulo R del plano uv con los lados paralelos a los ejes u y v , $R = \{(u, v) / u_1 \leq u \leq u_2, v_1 \leq v \leq v_2\}$, se transforma en un paralelogramo en el plano p con los lados paralelos a los vectores \vec{a} y \vec{b} .

El paralelogramo está generado por los vectores

$$\vec{r}(u_2, v_1) - \vec{r}(u_1, v_1) = (u_2\vec{a} + v_1\vec{b} + \vec{c}) - (u_1\vec{a} + v_1\vec{b} + \vec{c}) = (u_2 - u_1)\vec{a},$$

$$\vec{r}(u_1, v_2) - \vec{r}(u_1, v_1) = (u_1\vec{a} + v_2\vec{b} + \vec{c}) - (u_1\vec{a} + v_1\vec{b} + \vec{c}) = (v_2 - v_1)\vec{b}.$$



Luego el área del paralelogramo es

$$\| (u_2 - u_1)\vec{a} \times (v_2 - v_1)\vec{b} \| = \| \vec{a} \times \vec{b} \| (u_2 - u_1)(v_2 - v_1) = \| \vec{a} \times \vec{b} \| (\text{área de } R),$$

que siendo $\| \vec{a} \times \vec{b} \|$ constante, podemos escribir como

$$\text{área del paralelogramo} = \iint_R \| \vec{a} \times \vec{b} \| du dv.$$

Definición. Definimos el **área de superficie** $A(S)$ de una superficie parametrizada por

$$A(S) = \iint_D \| \vec{N}(u, v) \| du dv = \iint_D \| \vec{T}_u \times \vec{T}_v \| du dv, \quad (13.20)$$

donde $\| \vec{N}(u, v) \| = \| \vec{T}_u \times \vec{T}_v \|$ es la norma del vector producto vectorial fundamental. Si S es una unión de superficies S_i , su área es la suma de las áreas de las S_i .

Podemos justificar esta definición analizando la integral $\iint_D \| \vec{T}_u \times \vec{T}_v \| du dv$ en términos de sumas de Riemann. Por simplicidad, supongamos que D sea un rectángulo; consideremos

la n -ésima partición regular de D , y sea R_{ij} el ij -ésimo rectángulo en la partición, con vértices (u_i, v_i) , (u_{i+1}, v_i) , (u_i, v_{i+1}) y (u_{i+1}, v_{i+1}) , con $0 \leq i \leq n-1$, $0 \leq j \leq n-1$.

Si denotamos los vectores \vec{T}_u y \vec{T}_v en el punto (u_i, v_i) por \vec{T}_{u_i} y \vec{T}_{v_i} , podemos pensar en los vectores $\Delta u \vec{T}_{u_i}$ y $\Delta v \vec{T}_{v_i}$ como tangentes a la superficie en $\vec{\Phi}(u_i, v_i) = (x_{ij}, y_{ij}, z_{ij})$, donde $\Delta u = u_{i+1} - u_i$ y $\Delta v = v_{i+1} - v_i$. Entonces, estos vectores forman un paralelogramo P_{ij} que está en el plano tangente a la superficie en (x_{ij}, y_{ij}, z_{ij}) . Tenemos entonces una “cubierta de retazos” de la superficie mediante los paralelogramos P_{ij} . Para n grande, el área de P_{ij} es una buena aproximación al área de $\vec{\Phi}(R_{ij})$. Como el área del paralelogramo generado por dos vectores \vec{v}_1 y \vec{v}_2 es $\|\vec{v}_1 \times \vec{v}_2\|$, vemos que

$$A(P_{ij}) = \|\Delta u \vec{T}_{u_i} \times \Delta v \vec{T}_{v_i}\| = \|\vec{T}_{u_i} \times \vec{T}_{v_i}\| \Delta u \Delta v.$$

Por lo tanto, el área de la “cubierta de retazos” es

$$A_n = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} A(P_{ij}) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \|\vec{T}_{u_i} \times \vec{T}_{v_i}\| \Delta u \Delta v.$$

Cuando $n \rightarrow \infty$, las sumas A_n convergen a la integral

$$\iint_D \|\vec{T}_u \times \vec{T}_v\| du dv,$$

y como A_n deberá aproximar cada vez mejor el área de la superficie, la integral de arriba es una definición razonable de el área de la superficie $A(S)$.

Para estar seguros de que la fórmula (13.20) no contradice la noción de área establecida previamente, hemos de comprobar que nos da el resultado esperado tanto en el caso de regiones planas como en el caso de superficies de revolución.

Ejemplos.

- Como hemos visto en un ejemplo previo, la función

$$\vec{r}(u, v) = a \cos u \cos v \vec{i} + a \sin u \cos v \vec{j} + a \sin v \vec{k},$$

con $(u, v) \in \Omega$, siendo $\Omega = \{(u, v) / 0 \leq u \leq 2\pi, -\pi/2 \leq v \leq \pi/2\}$, parametriza una esfera de radio a centrada en el origen. Para esta parametrización tenemos

$$\vec{T}_u = -a \sin u \cos v \vec{i} + a \cos u \cos v \vec{j},$$

$$\vec{T}_v = -a \cos u \sin v \vec{i} + a \sin u \sin v \vec{j} + a \cos v \vec{k}.$$

Luego

$$\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = a \cos v \vec{r}(u, v),$$

es decir, como era de esperar el producto vectorial fundamental de una esfera es paralelo al radio vector $\vec{r}(u, v)$. Además,

$$\|\vec{N}(u, v)\| = \|\vec{T}_u \times \vec{T}_v\| = |a^2| |\cos v| = a^2 \cos v,$$

13.4. SUPERFICIES PARAMETRIZADAS.

por el hecho de que $-\pi/2 \leq v \leq \pi/2$.

Entonces, de acuerdo con la fórmula (13.20), tenemos

$$\begin{aligned} \text{área de la esfera} &= \iint_{\Omega} \|\vec{T}_u \times \vec{T}_v\| \, du \, dv = \iint_{\Omega} a^2 \cos v \, du \, dv = \\ &= \int_0^{2\pi} \left(\int_{-\pi/2}^{\pi/2} a^2 \cos v \, dv \right) du = 2\pi a^2 \int_{-\pi/2}^{\pi/2} \cos v \, dv = 4\pi a^2. \end{aligned}$$

- Sea S la superficie de una región plana Ω ; podemos parametrizar S haciendo

$$\vec{r}(u, v) = u\vec{i} + v\vec{j}, \quad (u, v) \in \Omega.$$

Aquí será

$$\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = \vec{i} \times \vec{j} = \vec{k}, \quad \text{luego } \|\vec{N}(u, v)\| = 1.$$

En este caso la fórmula (13.20) se reduce a la fórmula conocida

$$A = \iint_{\Omega} du \, dv.$$

- Sea ahora S generada al girar la gráfica de la función $y = f(x)$, con $x \in [a, b]$ alrededor del eje x . Supondremos que f es positiva y continuamente diferenciable.

Podemos parametrizar S haciendo

$$\vec{r}(u, v) = v\vec{i} + f(v) \cos u\vec{j} + f(v) \operatorname{sen} u\vec{k},$$

con $(u, v) \in \Omega$, siendo el conjunto $\Omega = \{(u, v) / 0 \leq u \leq 2\pi, a \leq v \leq b\}$.

Para esta parametrización tenemos

$$\begin{aligned} \vec{T}_u &= -f(v) \operatorname{sen} u\vec{j} + f(v) \cos u\vec{k}, \\ \vec{T}_v &= \vec{i} + f'(v) \cos u\vec{j} + f'(v) \operatorname{sen} u\vec{k}. \end{aligned}$$

Luego

$$\vec{N}(u, v) = \vec{T}_u \times \vec{T}_v = -f(v)f'(v)\vec{i} + f(v) \cos u\vec{j} + f(v) \operatorname{sen} u\vec{k},$$

con lo que

$$\|\vec{N}(u, v)\| = \|\vec{T}_u \times \vec{T}_v\| = f(v) \sqrt{[f'(v)]^2 + 1}.$$

Entonces, de acuerdo con la fórmula (13.20), tenemos

$$\begin{aligned} \text{área de } S &= \iint_{\Omega} \|\vec{T}_u \times \vec{T}_v\| \, du \, dv = \iint_{\Omega} f(v) \sqrt{[f'(v)]^2 + 1} \, du \, dv = \\ &= \int_0^{2\pi} \left(\int_a^b f(v) \sqrt{[f'(v)]^2 + 1} \, dv \right) du = 2\pi \int_a^b f(v) \sqrt{[f'(v)]^2 + 1} \, dv, \end{aligned}$$

que coincide con la fórmula dada en el subsección 3.6.4.

Dado que

$$\|\vec{T}_u \times \vec{T}_v\| = \sqrt{\left[\frac{\partial(x,y)}{\partial(u,v)}\right]^2 + \left[\frac{\partial(y,z)}{\partial(u,v)}\right]^2 + \left[\frac{\partial(x,z)}{\partial(u,v)}\right]^2}$$

donde

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}, \quad \frac{\partial(y,z)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{vmatrix}, \quad \frac{\partial(x,z)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{vmatrix},$$

podemos escribir la fórmula (13.20) como

$$A(S) = \iint_D \sqrt{\left[\frac{\partial(x,y)}{\partial(u,v)}\right]^2 + \left[\frac{\partial(y,z)}{\partial(u,v)}\right]^2 + \left[\frac{\partial(x,z)}{\partial(u,v)}\right]^2} du dv, \quad (13.21)$$

que a veces puede resultar más comoda de usar.

Ejemplo. Sea D la región determinada por $0 \leq \theta \leq 2\pi$ y $0 \leq r \leq 1$, y sea la función $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ definida por

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

una parametrización de un cono S . Halla su área de superficie.

Solución:

En la fórmula (13.21) tenemos

$$\begin{aligned} \frac{\partial(x,y)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r, \\ \frac{\partial(y,z)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin \theta & r \cos \theta \\ 1 & 0 \end{vmatrix} = -r \cos \theta, \\ \frac{\partial(x,z)}{\partial(r,\theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ 1 & 0 \end{vmatrix} = r \sin \theta, \end{aligned}$$

así que el integrando es

$$\|\vec{T}_u \times \vec{T}_v\| = \sqrt{r^2 + r^2 \cos^2 \theta + r^2 \sin^2 \theta} = r\sqrt{2}.$$

13.4. SUPERFICIES PARAMETRIZADAS.

Tenemos entonces,

$$A(S) = \int_0^{2\pi} \int_0^1 r \sqrt{2} dr d\theta = 2\sqrt{2}\pi \left. \frac{r^2}{2} \right|_0^1 = \sqrt{2}\pi.$$

Nótese que la norma del vector producto fundamental $\|\vec{T}_u \times \vec{T}_v\|$ es nula para $r = 0$, pero $\vec{\Phi}(0, \theta) = (0, 0, 0)$ para cualquier θ , así que $(0, 0, 0)$ no es el único punto donde la superficie no es suave.

Para confirmar que ésta es el área de $\vec{\Phi}(D)$ debemos verificar que $\vec{\Phi}$ es uno a uno para puntos que no estén en la frontera de D . Sea D° el conjunto de (r, θ) con $0 < r < 1$ y $0 < \theta < 2\pi$. Es decir, D° es el interior de D , esto es D sin su frontera. Para ver que $\vec{\Phi}$ es uno a uno, supongamos que $\vec{\Phi}(r, \theta) = \vec{\Phi}(r', \theta')$, para (r, θ) y $(r', \theta') \in D^\circ$. Entonces

$$r \cos \theta = r' \cos \theta', \quad r \sin \theta = r' \sin \theta', \quad r = r' \quad \Rightarrow \quad \theta = \theta' + 2k\pi.$$

Pero la última igualdad sólo es posible para $k = 0$, visto que (r, θ) y $(r', \theta') \in D^\circ$, esto es, tanto θ como θ' pertenecen al intervalo abierto $(0, 2\pi)$, y no pueden estar a 2π radianes de distancia. Esto prueba que fuera de la frontera, $\vec{\Phi}$ es uno a uno. En futuros ejemplos, por lo general no verificaremos que la parametrización sea uno a uno cuando resulte intuitivamente claro.

Si una superficie S viene dada en la forma $z = f(x, y)$, donde $(x, y) \in D$, hemos visto ya que ésta admite la parametrización

$$x = u, \quad y = v, \quad z = f(u, v), \quad \text{para } (u, v) \in D.$$

Cuando f es de clase C^1 , esta parametrización es suave, y la fórmula para el área de superficie se transforma en

$$A(S) = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} du dv, \quad (13.22)$$

habiendo aplicado las fórmulas

$$\vec{T}_u = \vec{i} + \frac{\partial f}{\partial u} \vec{k}, \quad \vec{T}_v = \vec{j} + \frac{\partial f}{\partial v} \vec{k}, \quad \vec{T}_u \times \vec{T}_v = -\frac{\partial f}{\partial u} \vec{i} - \frac{\partial f}{\partial v} \vec{j} + \vec{k}.$$

Ejemplo 1. Calcula el área de la parte del cilindro parabólico $z = y^2$ que está por encima del triángulo con vértices $(0, 0)$, $(0, 1)$ y $(1, 1)$ en el plano xy .

Solución:

Aquí, $f(x, y) = y^2$ y D es la región que se puede expresar como $0 \leq y \leq 1$, $0 \leq x \leq y$. Para aplicar la fórmula (13.22), identifiquemos $u = x$, $v = y$ y calculemos

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 2y,$$

luego

$$\begin{aligned} A(S^+) &= \iint_D \sqrt{1+4y^2} \, dx \, dy = \int_0^1 \int_0^y \sqrt{1+4y^2} \, dx \, dy = \\ &= \int_0^1 y \sqrt{1+4y^2} \, dy = \left[\frac{1}{12} (1+4y^2)^{3/2} \right]_0^1 = \frac{1}{12} (5\sqrt{5} - 1). \end{aligned}$$

Veamos ahora otro ejemplo en el cual además de aplicarse la fórmula (13.22), aparecen integrales impropias que trataremos exactamente como en el caso de funciones de una sola variable.

Ejemplo 2. Calcula el área de la superficie de la esfera S descrita por la ecuación $x^2 + y^2 + z^2 = 1$.

Solución:

Calcularemos el área del hemisfero superior S^+ dado por $x^2 + y^2 + z^2 = 1$ con $z \geq 0$, y luego multiplicaremos el resultado por 2. Por lo tanto tenemos,

$$z = f(x, y) = \sqrt{1 - x^2 - y^2}, \quad x^2 + y^2 \leq 1.$$

Sea D la región $x^2 + y^2 \leq 1$. Para aplicar la fórmula (13.22), calculamos

$$\frac{\partial f}{\partial x} = \frac{-x}{\sqrt{1-x^2-y^2}}, \quad \frac{\partial f}{\partial y} = \frac{-y}{\sqrt{1-x^2-y^2}},$$

luego

$$\begin{aligned} A(S^+) &= \iint_D \sqrt{1 + \left(\frac{-x}{\sqrt{1-x^2-y^2}} \right)^2 + \left(\frac{-y}{\sqrt{1-x^2-y^2}} \right)^2} \, du \, dv = \\ &= \iint_D \frac{1}{\sqrt{1-x^2-y^2}} \, du \, dv, \end{aligned}$$

que es una integral impropia. Sin embargo, podemos aplicar el teorema de Fubini en este caso para obtener la integral impropia iterada:

$$\begin{aligned} &= \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\sqrt{1-x^2-y^2}} \, dy \, dx = \int_{-1}^1 \left[\arcsen \frac{y}{\sqrt{1-x^2}} \right]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \, dx = \\ &= \int_{-1}^1 \left(\frac{\pi}{2} + \frac{\pi}{2} \right) \, dx = 2\pi. \end{aligned}$$

Así, el área de toda la esfera es 4π .

Ejercicios.

- 1) Deduce, usando la fórmula (13.21), la fórmula del área de la superficie lateral generada al girar la gráfica de una función $y = f(x)$ alrededor del eje x , con $x \in [a, b]$

$$A_x = 2\pi \int_a^b |f(v)| \sqrt{[f'(v)]^2 + 1} \, dv.$$

13.5. INTEGRALES DE FUNCIONES ESCALARES SOBRE SUPERFICIES.

- 2) Halla la superficie de la parte del paraboloido hiperbólico $z = xy$ que está dentro del cilindro $x^2 + y^2 = a^2$.
- 3) Halla el área de la parte del plano $bcx + acy + abz = abc$ que está dentro del primer octante ($x \geq 0, y \geq 0, z \geq 0$).
- 4) Halla el área de la superficie $z = x^2 + y^2$ desde $z = 0$ hasta $z = 4$.
- 5) Calcula el área de las siguientes superficies.
 - a) $z = a^2 - (x^2 + y^2)$ con $\frac{a^2}{4} \leq x^2 + y^2 \leq a^2$;
 - b) $3z = x^{3/2} + y^{3/2}$ con $0 \leq x \leq 1, 0 \leq y \leq x$;
 - c) $x^2 + y^2 + z^2 - 4z = 0$ con $0 \leq 3(x^2 + y^2) \leq z^2, z \geq 2$.
- 6) Sea la curva $y = f(x)$, con $x \in [a, b]$ alrededor del eje y . Demuestra que el área de la superficie barrida es

$$A_y = 2\pi \int_a^b |x| \sqrt{[f'(x)]^2 + 1} dx.$$

- 7) Halla el área de la superficie obtenida al girar la curva $y = x^2, 0 \leq x \leq 1$, alrededor del eje y .

13.5 Integrales de funciones escalares sobre superficies.

Ahora estamos preparados para definir la integral de una función escalar f sobre una superficie S , que es una generalización natural del área de una superficie, que corresponde a la integral sobre S de la función escalar $f(x, y, z) = 1$.

Comencemos con una superficie S parametrizada por una función $\vec{\Phi}: D \rightarrow S \subset \mathbb{R}^3$, $\vec{\Phi}(u, v) = (x(u, v), y(u, v), z(u, v))$.

Definición. Si $f(x, y, z)$ es una función continua con valores reales definida en S , definimos la **integral de f sobre S** como

$$\int_S f(x, y, z) dS = \int_S f dS = \int_D f(\vec{\Phi}(u, v)) \|\vec{T}_u \times \vec{T}_v\| du dv. \quad (13.23)$$

Desarrollando la ecuación (13.23) se transforma en

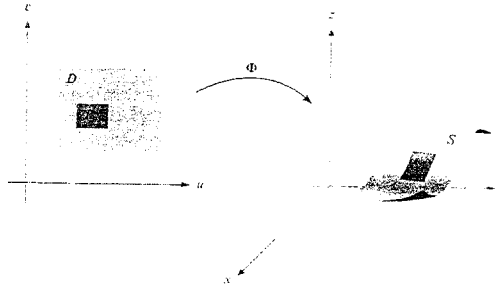
$$\int_S f dS = \int_D f(x(u, v), y(u, v), z(u, v)) \sqrt{\left[\frac{\partial(x, y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(y, z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(x, z)}{\partial(u, v)}\right]^2} du dv,$$

donde el radicando tiene el mismo significado que antes.

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Si $f(x, y, z)$ es idénticamente igual a 1, recuperamos la fórmula (13.21).

Podemos obtener un conocimiento intuitivo de esta integral al considerarla como límite de sumas. Sea D sea un rectángulo partido en n^2 rectángulos R_{ij} . Sea $S_{ij} = \vec{\Phi}(R_{ij})$ la parte de superficie $\vec{\Phi}(D)$ correspondiente a R_{ij} , y sea $A(S_{ij})$ el área de esta parte de superficie.



Para n grande, f será aproximadamente constante en S_{ij} ; construyamos la suma

$$S_n = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} f(\vec{\Phi}(u_i, v_j)) A(S_{ij}),$$

donde $(u_i, v_j) \in R_{ij}$. Usando la fórmula (13.20) y el teorema del valor medio para integrales, tenemos que

$$A(S_{ij}) = \iint_{R_{ij}} \|\vec{T}_u \times \vec{T}_v\| du dv = \|\vec{T}_{u_i^*} \times \vec{T}_{v_j^*}\| \Delta u \Delta v,$$

para algún $(u_i^*, v_j^*) \in R_{ij}$. Por lo tanto, la suma se convierte en

$$S_n = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} f(\vec{\Phi}(u_i, v_j)) \|\vec{T}_{u_i^*} \times \vec{T}_{v_j^*}\| \Delta u \Delta v,$$

que es una suma aproximante para la última integral en la fórmula (13.23). Por lo tanto

$$\lim_{n \rightarrow \infty} S_n = \int_S f dS.$$

Si S es unión de superficies parametrizadas $S_i, i = 1, \dots, N$, que no se intersectan excepto, quizá, a lo largo de curvas que definen sus fronteras, entonces la integral de f sobre S está definida por

$$\int_S f dS = \sum_{i=1}^N \int_{S_i} f dS.$$

Por ejemplo, como consecuencia de la fórmula anterior, la integral sobre la superficie de un cubo se puede expresar como la suma de las integrales sobre los seis lados.

13.5. INTEGRALES DE FUNCIONES ESCALARES SOBRE SUPERFICIES.

Ejemplo 1. Calcular la integral $\int_S f \, dS$ si la función $f(x, y, z)$ está dada por

$$f(x, y, z) = \sqrt{x^2 + y^2 + 1}$$

y la superficie S es la **helicoides** definida como $\vec{\Phi}: D \rightarrow \mathbb{R}^3$,

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

y D es la región donde $0 \leq \theta \leq 2\pi$ y $0 \leq r \leq 1$.

Solución:

Usaremos la fórmula (13.23). Por tanto, empezamos calculando

$$\begin{aligned} \frac{\partial(x, y)}{\partial(r, \theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r, \\ \frac{\partial(y, z)}{\partial(r, \theta)} &= \begin{vmatrix} \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin \theta & r \cos \theta \\ 0 & 1 \end{vmatrix} = \sin \theta, \\ \frac{\partial(x, z)}{\partial(r, \theta)} &= \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ 0 & 1 \end{vmatrix} = \cos \theta, \end{aligned}$$

así que

$$\|\vec{T}_r \times \vec{T}_\theta\| = \sqrt{r^2 + \sin^2 \theta + \cos^2 \theta} = \sqrt{r^2 + 1},$$

que nunca se anula, de modo que la superficie helicoides es suave.

Además, resulta

$$f(\vec{\Phi}(u, v)) = f(x(r, \theta), y(r, \theta), z(r, \theta)) = f(r \cos \theta, r \sin \theta, \theta) = \sqrt{r^2 + 1}.$$

Por lo tanto

$$\begin{aligned} \int_S f(x, y, z) \, dS &= \int_D f(\vec{\Phi}(r, \theta)) \|\vec{T}_r \times \vec{T}_\theta\| \, du \, dv = \\ &= \int_0^{2\pi} \int_0^1 \sqrt{r^2 + 1} \sqrt{r^2 + 1} \, dr \, d\theta = \int_0^{2\pi} \frac{4}{3} \, d\theta = \frac{8}{3} \pi. \end{aligned}$$

Ejemplo 2. Evalúa la integral $\int_S z^2 \, dS$, donde S es la esfera unitaria $x^2 + y^2 + z^2 = 1$.

Solución:

Para este problema es conveniente representar la esfera paramétricamente mediante las ecuaciones

$$x = \cos \theta \sin \phi, \quad y = \sin \theta \sin \phi, \quad z = \cos \phi,$$

sobre la región D en el plano $\theta\phi$ dada por $0 \leq \phi \leq \pi$ y $0 \leq \theta \leq 2\pi$.

Por un lado, haciendo los cálculos, resulta

$$\|\vec{T}_\theta \times \vec{T}_\phi\| = |\operatorname{sen} \phi|;$$

y por otro, de la fórmula (13.23) obtenemos

$$\begin{aligned} \int_S z^2 dS &= \int_D (\cos \phi)^2 \|\vec{T}_\theta \times \vec{T}_\phi\| d\theta d\phi = \\ &= \int_0^{2\pi} \int_0^\pi \cos^2 \phi |\operatorname{sen} \phi| d\phi d\theta = \int_0^{2\pi} \int_0^\pi \cos^2 \phi \operatorname{sen} \phi d\phi d\theta = \\ &= \frac{1}{3} \int_0^{2\pi} [-\cos^3 \phi]_0^\pi d\theta = \frac{2}{3} \int_0^{2\pi} d\theta = \frac{4}{3} \pi. \end{aligned}$$

Si S es la gráfica de una función C^1 , $z = g(x, y)$, vimos que podemos parametrizar la superficie S por

$$x = u, \quad y = v, \quad z = g(u, v).$$

En este caso

$$\|\vec{T}_u \times \vec{T}_v\| = \sqrt{1 + \left(\frac{\partial g}{\partial u}\right)^2 + \left(\frac{\partial g}{\partial v}\right)^2},$$

de modo que

$$\int_S f dS = \int_D f(x, y, g(x, y)) \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} dx dy. \quad (13.24)$$

Ejemplo. Calcula la integral $\int_S x dS$ si la superficie S viene definida por $z = x^2 + y$ y D es la región donde $0 \leq x \leq 1$ y $-1 \leq y \leq 1$.

Solución:

Usaremos la fórmula (13.24); empecemos poniendo

$$z = g(x, y) = x^2 + y$$

y calculando

$$\frac{\partial g}{\partial u} = \frac{\partial g}{\partial x} = 2x, \quad \frac{\partial g}{\partial v} = \frac{\partial g}{\partial y} = 1.$$

así que la fórmula (13.24) da

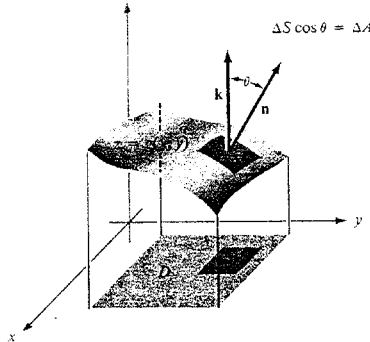
$$\begin{aligned} \int_S x dS &= \int_D f(x, y, g(x, y)) \sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} dx dy = \\ &= \int_{-1}^1 \int_0^1 x \sqrt{1 + 4x^2 + 1} dx dy = \sqrt{2} \left(\sqrt{3} - \frac{1}{3} \right). \end{aligned}$$

13.5. INTEGRALES DE FUNCIONES ESCALARES SOBRE SUPERFICIES.

Desarrollaremos ahora una fórmula para integrales de superficie cuando la superficie se puede representar como una gráfica. Para ello, sea S la gráfica de $z = g(x, y)$ y consideremos la fórmula (13.24). Al describir la superficie por la ecuación $\phi(x, y, z) = z - g(x, y) = 0$, un vector normal es

$$\vec{n} = \vec{\nabla}\phi = \left(-\frac{\partial g}{\partial x}, -\frac{\partial g}{\partial y}, 1 \right).$$

Consideremos ahora el vector unitario \vec{k} en el punto $(x, y, g(x, y))$ y sea θ el ángulo que forma la normal \vec{n} a la superficie con el vector \vec{k} .



Entonces

$$\cos \theta = \frac{\vec{n} \cdot \vec{k}}{\|\vec{n}\|} = \frac{1}{\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1}}.$$

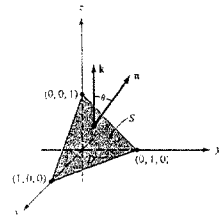
Al sustituir esta expresión en la fórmula (13.24) obtenemos

$$\int_S f(x, y, z) dS = \int_D \frac{f(x, y, g(x, y))}{\cos \theta} dx dy. \quad (13.25)$$

Ejemplo. Calcula la integral $\int_S x dS$ donde S es el triángulo con vértices $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$.

Solución:

Esta superficie es el plano descrito por la ecuación $x + y + z = 1$. Como la superficie es un plano, el ángulo θ es constante y un vector normal unitario es $\vec{n} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)$. Así, $\cos \theta = \vec{n} \cdot \vec{k} = \frac{1}{\sqrt{3}}$ y por la ecuación (13.25),



$$\int_S x dS = \sqrt{3} \int_D x dx dy = \sqrt{3} \int_0^1 \int_0^{1-x} x dy dx = \sqrt{3} \int_0^1 x(1-x) dx = \frac{\sqrt{3}}{6}.$$

Las integrales de funciones sobre superficies son útiles para calcular la masa de una superficie, cuando se conoce la función m de densidad de masa. La masa total de una superficie con densidad de masa, por unidad de área, $m(x, y, z)$ viene dada por

$$M(S) = \int_S m(x, y, z) dS.$$

Terminamos esta sección con

Algunas aplicaciones de las integrales de superficie.

Si S es una superficie de \mathbb{R}^3 que representa una lámina (no necesariamente plana) muy fina de densidad variable $f(x, y, z)$ en el punto $(x, y, z) \in S$ entonces:

a) La **masa** de la lámina es:

$$m(S) = \iint_S f dS. \tag{13.26}$$

b) Las **coordenadas del centro de gravedad** de la lámina. $(\bar{x}, \bar{y}, \bar{z})$, son:

$$\bar{x} = \frac{\iint_S x f(x, y, z) dS}{m(S)}, \quad \bar{y} = \frac{\iint_S y f(x, y, z) dS}{m(S)}, \quad \bar{z} = \frac{\iint_S z f(x, y, z) dS}{m(S)}. \tag{13.27}$$

c) Si L es una recta en \mathbb{R}^3 y $\delta(x, y, z)$ es la distancia del punto $(x, y, z) \in C$ a la recta L . el **momento de inercia** de la lámina respecto a L es:

$$I_L = \iint_S \delta(x, y, z)^2 f(x, y, z) dS.$$

En particular los momentos de inercia respecto a los ejes coordenados son:

$$\begin{aligned} I_x &= \iint_S (y^2 + z^2) f(x, y, z) dS \\ I_y &= \iint_S (x^2 + z^2) f(x, y, z) dS \\ I_z &= \iint_S (x^2 + y^2) f(x, y, z) dS. \end{aligned} \tag{13.28}$$

Ejercicios.

- 1) Halla el área de la superficie de la esfera S representada paramétricamente por $\vec{\Phi} : D \rightarrow S \subset \mathbb{R}^3$ donde D es el rectángulo $0 \leq \theta \leq 2\pi, 0 \leq \phi \leq \pi$ y $\vec{\Phi}$ está dada por las ecuaciones

$$x = \cos \theta \sin \phi, \quad y = \sin \theta \sin \phi, \quad z = \cos \phi.$$

13.6. INTEGRALES DE SUPERFICIE DE FUNCIONES VECTORIALES.

2) Halla el área de la superficie definida por $x + y + z = 1$, $x^2 + 2y^2 \leq 1$.

3) Calcula el área de la superficie dada por

$$x = r \cos \theta, \quad y = 2r \cos \theta, \quad z = \theta, \quad 0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi.$$

4) Calcula la integral de $x + y$ sobre la esfera unitaria.

5) Evalúa $\iint_S (x + y + z) dS$, donde S es la superficie esférica de radio unidad.

6) Encuentra la masa de una superficie esférica S de radio R tal que en cada punto $(x, y, z) \in S$ la densidad de masa es igual a la distancia de (x, y, z) hasta algún punto fijo $(x_0, y_0, z_0) \in S$.

7) Calcula las integrales siguientes sobre la superficie S : $z = \frac{1}{2}y^2$, con $0 \leq x \leq 1$ y $0 \leq y \leq 1$.

a) $\iint_S dS$;

b) $\iint_S x^2 dS$;

c) $\iint_S 3y dS$;

d) $\iint_S (x - y) dS$;

e) $\iint_S \sqrt{2z} dS$;

f) $\iint_S \sqrt{1 + y^2} dS$.

8) Halla la masa de una superficie material con forma del triángulo $(a, 0, 0)$, $(0, a, 0)$, $(0, 0, a)$, sabiendo que la función de densidad es:

a) $m(x, y, z) = k$;

b) $m(x, y, z) = k(x + y)$;

c) $m(x, y, z) = kx^2$.

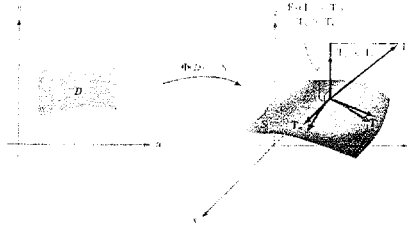
13.6 Integrales de superficie de funciones vectoriales.

En esta sección nos ocuparemos de integrales de funciones vectoriales sobre superficies. La definición que daremos es una extensión natural de la definición dada para funciones escalares estudiadas en la sección 13.5.

Definición. Sea $\vec{F}(x, y, z)$ un campo vectorial definido en S , imagen de una superficie parametrizada $\vec{\Phi}$. Definimos la **integral de superficie de \vec{F} sobre $\vec{\Phi}$** como

$$\int_{\vec{\Phi}} \vec{F}(x, y, z) \cdot d\vec{S} = \int_D \vec{F}(\vec{\Phi}(u, v)) \cdot (\vec{T}_u \times \vec{T}_v) du dv, \quad (13.29)$$

donde \vec{T}_u y \vec{T}_v se definen como en la sección 13.4.



Ejemplo. Sea D el rectángulo en el plano $\theta\phi$ definido por

$$0 \leq \theta \leq 2\pi, \quad 0 \leq \phi \leq \pi.$$

y sea la superficie S definida por la parametrización $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ dada por

$$x = \cos \theta \sin \phi, \quad y = \sin \theta \sin \phi, \quad z = \cos \phi.$$

Sea \vec{r} el vector posición $\vec{r}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$. Calcula $\int_{\vec{\Phi}} \vec{r}(x, y, z) \cdot d\vec{S}$.

Solución:

Claramente, en este ejemplo, θ y ϕ son los ángulos de coordenadas esféricas, y S es la esfera unitaria parametrizada por $\vec{\Phi}$.

Primero hallamos

$$\begin{aligned} \vec{T}_\theta &= (-\sin \phi \sin \theta) \vec{i} + (\sin \phi \cos \theta) \vec{j} \\ \vec{T}_\phi &= (\cos \phi \cos \theta) \vec{i} + (\cos \phi \sin \theta) \vec{j} - (\sin \phi) \vec{k} \end{aligned}$$

entonces

$$\vec{T}_\theta \times \vec{T}_\phi = (-\sin^2 \phi \cos \theta) \vec{i} - (\sin^2 \phi \sin \theta) \vec{j} - (\sin \phi \cos \phi) \vec{k}.$$

Ahora calculamos

$$\begin{aligned} \vec{r} \cdot (\vec{T}_\theta \times \vec{T}_\phi) &= (x\vec{i} + y\vec{j} + z\vec{k}) \cdot (\vec{T}_\theta \times \vec{T}_\phi) = \\ &= [(\cos \theta \sin \phi) \vec{i} + (\sin \theta \sin \phi) \vec{j} + (\cos \phi) \vec{k}] \cdot \\ &\quad \cdot (-\sin \phi) [(\sin \phi \cos \theta) \vec{i} + (\sin \phi \sin \theta) \vec{j} + (\cos \phi) \vec{k}] = \\ &= (-\sin \phi) [\sin^2 \phi \cos^2 \theta + \sin^2 \phi \sin^2 \theta + \cos^2 \phi] = \\ &= -\sin \phi. \end{aligned}$$

Y finalmente

$$\int_{\vec{\Phi}} \vec{r}(x, y, z) \cdot d\vec{S} = \int_D -\sin \phi \, d\phi \, d\theta = \int_0^{2\pi} (-2) \, d\theta = -4\pi.$$

Podemos encontrar una analogía entre la integral de superficie $\int_{\vec{\Phi}} \vec{F}(x, y, z) \cdot d\vec{S}$ y la integral de línea $\int_{\vec{\sigma}} \vec{F}(x, y, z) \cdot d\vec{s}$, que es una integral orientada: hemos necesitado del

13.6. INTEGRALES DE SUPERFICIE DE FUNCIONES VECTORIALES.

concepto de orientación de una curva para extender la definición de $\int_{\vec{\sigma}} \vec{F}(x, y, z) \cdot d\vec{s}$ a integrales de línea $\int_C \vec{F}(x, y, z) \cdot d\vec{s}$ sobre curvas orientadas. Extendemos la definición de $\int_{\vec{\Phi}} \vec{F}(x, y, z) \cdot d\vec{S}$ a superficies orientadas de manera similar; esto es, dada una superficie S parametrizada por una función $\vec{\Phi}$, queremos definir $\int_S \vec{F}(x, y, z) \cdot d\vec{S} = \int_{\vec{\Phi}} \vec{F}(x, y, z) \cdot d\vec{S}$ y mostrar que es independiente de la parametrización, excepto, quizá, por el signo. Para lograrlo necesitamos el concepto de orientación de una superficie.

Definición. Una **superficie orientada** S es una superficie de dos lados, uno de ellos llamado el **lado exterior** o **positivo**, y el otro el **lado interior** o **negativo**. En cada punto $(x, y, z) \in S$ hay dos vectores normales unitarios \vec{n}_1 y \vec{n}_2 , donde $\vec{n}_1 = -\vec{n}_2$. Cada una de estas normales se puede asociar con un lado de la superficie. Así, para especificar un lado de una superficie S , en cada punto escogemos un vector normal unitario \vec{n} que apunta hacia afuera desde el lado positivo de S en este punto.

Esta definición supone que nuestra superficie tiene dos lados. Hay ejemplos de superficies con un solo lado, el más famoso es la cinta de Möbius. La selección del lado llamado exterior, a menudo es impuesta por la superficie misma, como, por ejemplo, en el caso de una esfera. En otros casos, la denominación es algo arbitraria.

Sea $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ una parametrización de una superficie orientada S y supongamos que S es suave en $\vec{\Phi}(u_0, v_0)$, con $(u_0, v_0) \in D$; esto es, está definido el vector normal unitario

$$\frac{\vec{T}_{u_0} \times \vec{T}_{v_0}}{\|\vec{T}_{u_0} \times \vec{T}_{v_0}\|}.$$

Si $\vec{n}(\vec{\Phi}(u_0, v_0))$ denota la normal unitaria a S en $\vec{\Phi}(u_0, v_0)$ apuntando desde el lado positivo de S a ese punto, se sigue que $\frac{\vec{T}_{u_0} \times \vec{T}_{v_0}}{\|\vec{T}_{u_0} \times \vec{T}_{v_0}\|} = \pm \vec{n}(\vec{\Phi}(u_0, v_0))$.

Se dice que la parametrización $\vec{\Phi}$ **preserva la orientación** si $\frac{\vec{T}_u \times \vec{T}_v}{\|\vec{T}_u \times \vec{T}_v\|} = \vec{n}(\vec{\Phi}(u, v))$ en todo $(u, v) \in D$ para los cuales S es suave en $\vec{\Phi}(u, v)$. En otras palabras, $\vec{\Phi}$ preserva la orientación si el vector $\vec{T}_u \times \vec{T}_v$ apunta hacia afuera desde el lado exterior de la superficie.

Al contrario, se dice que $\vec{\Phi}$ **invierte la orientación** si $\frac{\vec{T}_u \times \vec{T}_v}{\|\vec{T}_u \times \vec{T}_v\|} = -\vec{n}(\vec{\Phi}(u, v))$ en todo $(u, v) \in D$ para los cuales S es suave en $\vec{\Phi}(u, v)$.

Ejemplo. Podemos dar a la esfera unitaria $x^2 + y^2 + z^2 = 1$ en \mathbb{R}^3 una orientación seleccionando el vector unitario $\vec{n}(x, y, z) = \vec{r}$, donde $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$, que apunta hacia afuera desde el lado exterior de la superficie (esta selección corresponde a nuestra concepción intuitiva del exterior de la esfera).

Volvamos a considerar la parametrización del ejemplo anterior: $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ dada por

$$x = \cos \theta \sin \phi, \quad y = \sin \theta \sin \phi, \quad z = \cos \phi.$$

El producto vectorial de los vectores tangentes \vec{T}_θ y \vec{T}_ϕ , esto es, una normal a S , está dado por

$$\vec{T}_\theta \times \vec{T}_\phi = (-\operatorname{sen} \phi) \left[(\operatorname{sen} \phi \cos \theta) \vec{i} + (\operatorname{sen} \phi \operatorname{sen} \theta) \vec{j} + (\cos \phi) \vec{k} \right] = -\vec{r} \operatorname{sen} \phi.$$

Como $-\operatorname{sen} \phi \leq 0$ para $0 \leq \phi \leq \pi$, este vector normal apunta hacia adentro desde la esfera. Así, la parametrización $\vec{\Phi}$ dada invierte la orientación.

Sea ahora S una superficie descrita por $z = f(x, y)$. Hay dos vectores normales unitarios a S en $(x_0, y_0, f(x_0, y_0))$, esto es $\pm \vec{n}$ donde

$$\vec{n} = \frac{-\frac{\partial f}{\partial x}(x_0, y_0) \vec{i} - \frac{\partial f}{\partial y}(x_0, y_0) \vec{j} + \vec{k}}{\sqrt{\left[\frac{\partial f}{\partial x}(x_0, y_0)\right]^2 + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^2 + 1}}.$$

Podemos orientar todas estas superficies tomando el lado positivo de S como el lado desde el cual apunta \vec{n} ; esto es, el lado positivo de dicha superficie está determinado por la normal unitaria \vec{n} con componente k positiva. Si parametrizamos esta superficie por $\vec{\Phi}(u, v) = (u, v, f(u, v))$, entonces $\vec{\Phi}$ preservará la orientación.

Teorema 13.9 *Sea \vec{F} un campo vectorial continuo definido en la superficie orientada S , y sean $\vec{\Phi}_1$ y $\vec{\Phi}_2$ dos parametrizaciones, de manera que la primera preserva la orientación. Entonces*

$$\int_{\vec{\Phi}_1} \vec{F} \cdot d\vec{S} = \begin{cases} \int_{\vec{\Phi}_2} \vec{F} \cdot d\vec{S} & \text{si } \vec{\Phi}_2 \text{ preserva la orientación,} \\ -\int_{\vec{\Phi}_2} \vec{F} \cdot d\vec{S} & \text{si } \vec{\Phi}_2 \text{ invierte la orientación.} \end{cases} \quad (13.30)$$

Si f es una función continua con valores reales definida en S , y si $\vec{\Phi}_1$ y $\vec{\Phi}_2$ son dos parametrizaciones de S , entonces

$$\int_{\vec{\Phi}_1} f \, dS = \int_{\vec{\Phi}_2} f \, dS. \quad (13.31)$$

Si $f = 1$, obtenemos

$$A(S) = \int_{\vec{\Phi}_1} dS = \int_{\vec{\Phi}_2} dS.$$

demostrando así que el área es independiente de la parametrización.

La demostración de este teorema es similar a la del teorema 13.3.

Podemos usar entonces sin ambigüedad la notación

$$\int_{\vec{\Phi}} \vec{F} \cdot d\vec{S} = \int_S \vec{F} \cdot d\vec{S},$$

13.6. INTEGRALES DE SUPERFICIE DE FUNCIONES VECTORIALES.

donde $\vec{\Phi}$ es una parametrización que preserva la orientación.

Otra analogía con las integrales de línea es la siguiente: vimos que una integral de línea $\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s}$ se puede pensar como la integral de trayectoria de la componente tangencial de \vec{F} a lo largo de $\vec{\sigma}$. Ahora, para una superficie suave orientada S y una parametrización $\vec{\Phi}$ que preserva la orientación de S , podemos expresar $\int_S \vec{F} \cdot d\vec{S}$ como integral de una función f con valores reales, sobre la superficie. Sea $\vec{n} = \frac{\vec{T}_u \times \vec{T}_v}{\|\vec{T}_u \times \vec{T}_v\|}$ la normal unitaria que apunta al exterior de S . Entonces

$$\begin{aligned} \int_S \vec{F} \cdot d\vec{S} &= \int_{\vec{\Phi}} \vec{F} \cdot d\vec{S} = \int_D \vec{F}(\vec{\Phi}(u, v)) \cdot (\vec{T}_u \times \vec{T}_v) \, du \, dv = \\ &= \int_D \vec{F}(\vec{\Phi}(u, v)) \cdot \left(\frac{\vec{T}_u \times \vec{T}_v}{\|\vec{T}_u \times \vec{T}_v\|} \right) \|\vec{T}_u \times \vec{T}_v\| \, du \, dv = \\ &= \int_D \vec{F} \cdot \vec{n} \|\vec{T}_u \times \vec{T}_v\| \, du \, dv = \int_S (\vec{F} \cdot \vec{n}) \, dS = \int_S f \, dS, \end{aligned}$$

donde $f = \vec{F} \cdot \vec{n}$. Así, hemos probado el siguiente teorema.

Teorema 13.10 $\int_S \vec{F} \cdot d\vec{S}$, la integral de superficie del campo vectorial \vec{F} sobre S , es igual a la integral de la componente normal de \vec{F} sobre la superficie; esto es

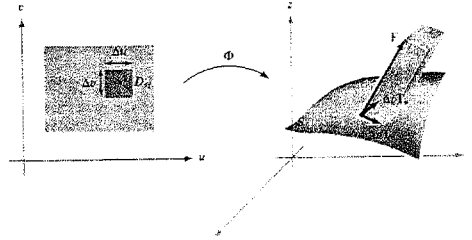
$$\int_S \vec{F} \cdot d\vec{S} = \int_S (\vec{F} \cdot \vec{n}) \, dS. \quad (13.32)$$

El significado geométrico y físico de la integral de superficie se puede entender expresándole como un límite de sumas de Riemann. Por sencillez, supongamos que D es un rectángulo. Fijemos una parametrización $\vec{\Phi}$ de S que preserve la orientación y partamos la región D en n^2 piezas D_{ij} , $0 \leq i \leq n-1$, $0 \leq j \leq n-1$. Denotemos por Δu la longitud del lado horizontal de D_{ij} y por Δv la longitud del lado vertical de D_{ij} . Sean (u, v) un punto en D_{ij} y $(x, y, z) = \vec{\Phi}(u, v)$, el punto correspondiente en la superficie. Consideremos el paralelogramo con lados $\Delta u \vec{T}_u$ y $\Delta v \vec{T}_v$ que está en el plano tangente a S en (x, y, z) y el paralelepípedo formado por \vec{F} , $\Delta u \vec{T}_u$ y $\Delta v \vec{T}_v$. El volumen del paralelepípedo es el valor absoluto del triple producto

$$\vec{F} \cdot (\Delta u \vec{T}_u \times \Delta v \vec{T}_v) = \vec{F} \cdot (\vec{T}_u \times \vec{T}_v) \Delta u \Delta v.$$

El vector $\vec{T}_u \times \vec{T}_v$ es normal a la superficie en (x, y, z) y apunta hacia afuera desde el exterior de la superficie. Así, el número $\vec{F} \cdot (\vec{T}_u \times \vec{T}_v)$ es positivo cuando el paralelepípedo está en el exterior de la superficie. En general, el paralelepípedo está en aquel lado de la superficie desde donde apunta \vec{F} . Si pensamos \vec{F} como el campo de velocidad de un fluido,

$\vec{F}(x, y, z)$ apunta en la dirección en la cual el fluido se mueve a través de la superficie cerca de (x, y, z) .



Más aún, el número

$$\left| \vec{F} \cdot \left(\Delta u \vec{T}_u \times \Delta v \vec{T}_v \right) \right|$$

mide la cantidad de fluido que pasa a través del paralelogramo tangente por unidad de tiempo. Como el signo de $\vec{F} \cdot \left(\Delta u \vec{T}_u \times \Delta v \vec{T}_v \right)$ es positivo si el vector \vec{F} apunta hacia afuera en (x, y, z) y negativo si \vec{F} apunta hacia adentro, $\sum_{ij} \vec{F} \cdot \left(\Delta u \vec{T}_u \times \Delta v \vec{T}_v \right)$ es una medida aproximada de la cantidad neta de fluido que fluye hacia afuera a través de la superficie por unidad de tiempo (recordar que “afuera” o “adentro” depende de la parametrización escogida). Por lo tanto, la integral $\int_S \vec{F} \cdot d\vec{S}$ es la cantidad neta de fluido que fluye a través de la superficie por unidad de tiempo, esto es, la tasa de flujo. Por lo tanto, esta integral también se llama flujo de \vec{F} a través de la superficie.

En el caso en que \vec{F} representa un campo magnético o eléctrico, $\int_S \vec{F} \cdot d\vec{S}$ también se conoce como flujo. El lector quizá conozca las leyes físicas (como la de Faraday) que relaciona el flujo de un campo vectorial con la circulación (o corriente) en un lazo circundante. Ésta es la base histórica y física del teorema de Stokes, que estudiaremos en la sección 13.8.

Las integrales de superficie también se aplican al estudio de flujo de calor. Sea $T(x, y, z)$ la temperatura en un punto $(x, y, z) \in W \subset \mathbb{R}^3$, donde W es alguna región y T es una función \mathcal{C}^1 . Entonces

$$\vec{\nabla}T = \frac{\partial T}{\partial x} \vec{i} + \frac{\partial T}{\partial y} \vec{j} + \frac{\partial T}{\partial z} \vec{k}$$

representa el gradiente de temperatura, y el calor “fluye” según el campo vectorial

$$-k\vec{\nabla}T = \vec{F},$$

donde k es una constante positiva. Por lo tanto $\int_S \vec{F} \cdot d\vec{S}$ es la tasa total de flujo de calor o flujo a través de la superficie S .

Ejemplo 1. Supongamos que una función temperatura está dada por $T(x, y, z) = x^2 + y^2 + z^2$ y sea S la esfera unitaria $x^2 + y^2 + z^2 = 1$ orientada con la normal exterior. Halla el flujo de calor a través de la superficie S si $k = 1$.

13.6. INTEGRALES DE SUPERFICIE DE FUNCIONES VECTORIALES.

Solución:

Tenemos

$$\vec{F} = \vec{\nabla}T(x, y, z) = -2x\vec{i} - 2y\vec{j} - 2z\vec{k}.$$

En S , $\vec{n}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$ es la normal "exterior" unitaria a S en (x, y, z) y $f(x, y, z) = \vec{F} \cdot \vec{n} = -2x^2 - 2y^2 - 2z^2 = -2$ es la componente normal de \vec{F} . En el teorema 13.10 podemos ver que la integral superficie de \vec{F} es igual a la integral de su componente normal $f = \vec{F} \cdot \vec{n}$ sobre S . Así,

$$\int_S \vec{F} \cdot d\vec{S} = \int_S f dS = \int_S (\vec{F} \cdot \vec{n}) dS = -2 \int_S dS = -2 A(S) = -2(4\pi) = -8\pi.$$

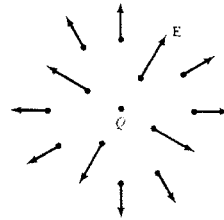
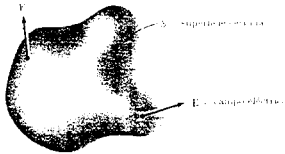
El flujo de calor se dirige hacia el centro de la esfera (¿por qué hacia el centro?). Está claro que nuestra observación de que $\int_S \vec{F} \cdot d\vec{S} = \int_S f dS$ nos ha ahorrado considerable tiempo de computación. En este ejemplo, $\vec{F}(x, y, z) = -2x\vec{i} - 2y\vec{j} - 2z\vec{k}$ podría también representar un campo eléctrico, en cuyo caso $\int_S \vec{F} \cdot d\vec{S} = -8\pi$ sería el flujo eléctrico a través de S .

Ejemplo 2. Hay una importante ley física, debida al gran matemático y físico K. F. Gauss, que relaciona el flujo de un campo eléctrico E a través de una superficie "cerrada" S (por ejemplo, una esfera o un elipsoide) con la carga neta Q encerrada por la superficie, a saber,

$$\int_E \vec{E} \cdot d\vec{S} = \frac{Q}{\varepsilon},$$

donde ε es la permitibilidad del medio donde se encuentra colocada dicha carga.

La ley de Gauss se estudiará en detalle en la sección 13.9.



Supongamos que $\vec{E} = E\vec{n}$; esto es, \vec{E} es múltiplo escalar constante de la normal unitaria a S . Entonces la ley de Gauss viene dada por

$$\int_E \vec{E} \cdot d\vec{S} = \int_S E dS = E \int_S dS = \frac{Q}{\varepsilon}.$$

Así,

$$E = \frac{Q}{A(S)\varepsilon}.$$

En el caso de que S sea la esfera de radio R , esta última ecuación se convierte en

$$E = \frac{Q}{A(S)\varepsilon} = \frac{Q}{4\pi R^2\varepsilon}.$$

Supongamos ahora que E está generado por una carga puntual aislada Q . Por simetría, se sigue que $\vec{E} = E\vec{n}$, donde \vec{n} es la normal unitaria a cualquier esfera con centro en Q . Por lo tanto se cumple la ecuación

$$E = \frac{Q}{4\pi R^2\varepsilon}.$$

Consideremos una segunda carga puntual Q_0 situada a una distancia R de Q . La fuerza \vec{F} que actúa sobre esta segunda carga Q_0 está dada por

$$\vec{F} = \vec{E}Q_0 = EQ_0\vec{n} = \frac{QQ_0}{4\pi R^2\varepsilon}\vec{n}.$$

Si F es la magnitud de \vec{F} , tenemos

$$F = \frac{QQ_0}{4\pi R^2\varepsilon}$$

que es la conocida ley de Coulomb para la fuerza entre dos cargas puntuales.

Para concluir esta sección deduzcamos las fórmulas de integrales de superficie para gráficas de funciones. Sea S la superficie descrita por $z = f(x, y)$ con $(x, y) \in D$, y orientada de modo que el vector unitario normal

$$\vec{n} = \frac{-\frac{\partial f}{\partial x}\vec{i} - \frac{\partial f}{\partial y}\vec{j} + \vec{k}}{\sqrt{\left[\frac{\partial f}{\partial x}\right]^2 + \left[\frac{\partial f}{\partial y}\right]^2 + 1}}$$

apunte hacia afuera.

Podemos parametrizar S por $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ dada por $\vec{\Phi}(x, y) = (x, y, f(x, y))$. En este caso la integral $\int_S \vec{F} \cdot d\vec{S}$ se puede escribir de manera más simple. Tenemos

$$\vec{T}_x = \vec{i} + \frac{\partial f}{\partial x}\vec{k}, \quad \vec{T}_y = \vec{j} + \frac{\partial f}{\partial y}\vec{k}.$$

Así,

$$\vec{T}_x \times \vec{T}_y = -\frac{\partial f}{\partial x}\vec{i} - \frac{\partial f}{\partial y}\vec{j} + \vec{k}.$$

Si $\vec{F} = (F_1, F_2, F_3)$ es un campo vectorial continuo, obtenemos la fórmula

$$\begin{aligned} \int_S \vec{F} \cdot d\vec{S} &= \int_D \vec{F} \cdot (\vec{T}_x \times \vec{T}_y) \, dx \, dy = \\ &= \int_D \left[F_1 \left(-\frac{\partial f}{\partial x} \right) + F_2 \left(-\frac{\partial f}{\partial y} \right) + F_3 \right] \, dx \, dy. \end{aligned} \tag{13.33}$$

13.6. INTEGRALES DE SUPERFICIE DE FUNCIONES VECTORIALES.

Ejemplo. Las ecuaciones

$$z = 12, \quad x^2 + y^2 \leq 25,$$

describen un disco de radio 5 que está en el plano $z = 12$. Supongamos que \vec{r} es el campo vectorial

$$\vec{r}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}.$$

Calcula $\int_S \vec{r} \cdot d\vec{S}$.

Solución:

Lo haremos de tres maneras. Primero, tenemos

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} = 0,$$

pues $z = 12$ es constante en el disco, de modo que

$$\vec{r}(x, y, z) \cdot (\vec{T}_x \times \vec{T}_y) = \vec{r}(x, y, z) \cdot (\vec{i} \times \vec{j}) = \vec{r}(x, y, z) \cdot \vec{k} = z$$

y entonces, usando la definición original dada al principio de esta sección, la integral se convierte en

$$\int_S \vec{r} \cdot d\vec{S} = \int_D z \, dx \, dy = \int_D 12 \, dx \, dy = 12 (\text{área de } D) = 12 (25\pi) = 300\pi.$$

Una segunda solución: como el disco es paralelo al plano xy , la normal unitaria exterior es \vec{k} . Entonces $\vec{n}(x, y, z) = \vec{k}$ y $\vec{r} \cdot \vec{n} = z$. Además, $\|\vec{T}_x \times \vec{T}_y\| = \|\vec{k}\| = 1$, y del teorema 13.10 se sigue que

$$\int_S \vec{r} \cdot d\vec{S} = \int_S \vec{r} \cdot \vec{n} \, dS = \int_S z \, dS = \int_D 12 \, dx \, dy = 300\pi.$$

Finalmente la tercera manera es utilizando la fórmula (13.33), con $f(x, y) = 12$ y D el disco $x^2 + y^2 = 25$:

$$\int_S \vec{r} \cdot d\vec{S} = \int_D (x \cdot 0 + y \cdot 0 + 12) \, dx \, dy = 12 (\text{área de } D) = 300\pi.$$

Ejercicios.

- 1) Sea $\vec{F}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k}$. Evalúa $\int_S \vec{F} \cdot d\vec{S}$, donde S es el hemisferio superior de la esfera unitaria $x^2 + y^2 + z^2 = 1$.
- 2) Sea $\vec{F} = x^2\vec{i} + y^2\vec{j} + z^2\vec{k}$. Evalúa $\int_S \vec{F} \cdot d\vec{S}$, donde S es la cara exterior de la semiesfera $x^2 + y^2 + z^2 = a^2$, con $z \geq 0$.

- 3) Evalúa $\iint_S \nabla \times \vec{F} \cdot d\vec{S}$, donde S es la superficie dada por $x^2 + y^2 + 3z^2 = 1, z \leq 0$ y $\vec{F}(x, y, z) = y\vec{i} - x\vec{j} + zx^3y^2\vec{k}$.
- 4) Sea el campo de velocidad de un fluido (expresado en m/seg) descrito por $\vec{F} = \vec{i} + x\vec{j} + z\vec{k}$. Calcula cuántos metros cúbicos de fluido cruzan por segundo la superficie $x^2 + y^2 + z^2 = 1, z \geq 0$.
- 5) Sea la temperatura de un punto en \mathbb{R}^3 dada por $T(x, y, z) = 3x^2 + 3z^2$. Calcula el flujo de calor a través de la superficie $x^2 + z^2 = 2, 0 \leq y \leq 2$, si $k = 1$.
- 6) Determina el flujo a través de $S: Z = xy$ con $0 \leq x \leq 1, 0 \leq y \leq 2$ en la dirección de la normal unitaria superior de los campos vectoriales siguientes.
- a) $\vec{F} = -xy^2\vec{i} + z\vec{j}$; b) $\vec{F} = xz\vec{j} - xy\vec{k}$; c) $\vec{F} = x^2y\vec{i} - z^2\vec{k}$.

13.7 Teorema de Green.

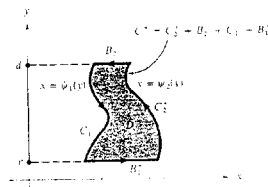
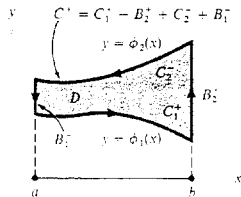
El teorema de Green relaciona una integral de línea a lo largo de una curva cerrada C en el plano \mathbb{R}^2 , con una integral doble sobre la región encerrada por C .

Este importante resultado será generalizado en las siguientes secciones, a curvas y superficies en \mathbb{R}^2 . Nos referiremos a integrales de línea alrededor de curvas que son fronteras de regiones elementales del tipo 1, 2 ó 3 (ver la sección 12.3).

Una curva cerrada simple C que es la frontera de una región del tipo 1, 2 ó 3 tiene dos orientaciones -en sentido contrario al que giran las manecillas del reloj (positiva) y en el sentido en que giran las manecillas del reloj (negativa)-. Denotamos C con la orientación en sentido contrario al que giran las manecillas del reloj por C^+ , y con la orientación en el sentido en que giran las manecillas del reloj por C^- . La frontera C de una región del tipo 1 se puede descomponer en partes superior e inferior, C_1 y C_2 , y (si es posible) partes verticales izquierda y derecha, B_1 y B_2 . Entonces escribimos,

$$C^+ = C_1^+ + B_2^+ + C_2^- + B_1^-$$

donde los signos de suma denotan las curvas orientadas en la dirección izquierda a derecha o de abajo hacia arriba, y los signos de resta denotan las curvas orientadas de derecha a izquierda o de arriba hacia abajo.



13.7. TEOREMA DE GREEN.

Podemos hacer descomposiciones similares de la frontera de una región de tipo 2 en partes izquierda y derecha, y partes horizontales superior e inferior (si es posible).

De manera análoga, la frontera de una región del tipo 3 tiene dos descomposiciones -una en mitades superior e inferior, la otra en mitades izquierda y derecha.

Teorema 13.11 Teorema de Green.

Sea D una región del tipo 3 y sea C su frontera. Supongamos que $P: D \rightarrow \mathbb{R}$ y $Q: D \rightarrow \mathbb{R}$ sean de clase C^1 . Entonces

$$\int_{C^+} P dx + Q dy = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy. \quad (13.34)$$

Demostración:

Dividiremos la demostración en dos partes: primero veremos que

$$\int_{C^+} P dx = - \int_D \frac{\partial P}{\partial y} dx dy,$$

y luego

$$\int_{C^+} Q dy = \int_D \frac{\partial Q}{\partial x} dx dy.$$

Supongamos que la región D está descrita por

$$a \leq x \leq b, \quad \phi_1(x) \leq y \leq \phi_2(x).$$

Descomponemos C^- escribiendo $C^+ = C_1^+ + B_2^+ + C_2^- + B_2^-$. Por el teorema de Fubini podemos evaluar la integral doble como una integral iterada y después usar el teorema fundamental de cálculo:

$$\int_D \frac{\partial P}{\partial y} dx dy = \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} \frac{\partial P}{\partial y} dy dx = \int_a^b [P(x, \phi_2(x)) - P(x, \phi_1(x))] dx.$$

Sin embargo, como C_1^+ se puede parametrizar por $x \rightsquigarrow (x, \phi_1(x))$, $a \leq x \leq b$, y C_2^+ se puede parametrizar por $x \rightsquigarrow (x, \phi_2(x))$, $a \leq x \leq b$, tenemos

$$\int_a^b P(x, \phi_1(x)) dx = \int_{C_1^+} P(x, y) dx \quad \text{y} \quad \int_a^b P(x, \phi_2(x)) dx = \int_{C_2^+} P(x, y) dx.$$

Así, al invertir orientaciones,

$$- \int_a^b P(x, \phi_2(x)) dx = \int_{C_2^-} P(x, y) dx.$$

Por lo tanto

$$\int_D \frac{\partial P}{\partial y} dx dy = - \int_{C_1^+} P(x, y) dx - \int_{C_2^-} P(x, y) dx$$

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Como x es constante en B_2^+ y B_1^- , tenemos

$$\int_{B_2^+} P dx = 0 = \int_{B_1^-} P dx$$

de modo que

$$\int_{C^+} P dx = \int_{C_1^+} P dx + \int_{B_2^+} P dx + \int_{C_2^-} P dx + \int_{B_1^-} P dx = \int_{C_1^+} P dx + \int_{C_2^-} P dx.$$

Así,

$$\int_D \frac{\partial P}{\partial y} dx dy = - \int_{C_1^+} P dx - \int_{C_2^-} P dx = - \int_{C^+} P dx.$$

Probaremos ahora la otra parte intercambiando los papeles de x e y . Supongamos que D está dada por

$$\psi_1(y) \leq y \leq \psi_2(y), \quad c \leq y \leq d.$$

Entonces podemos escribir $C^+ = C_1^- + B_1^+ + C_2^+ + B_2^-$, donde C_2^+ es la curva parametrizada por $y \rightsquigarrow (\psi_2(y), y)$, $c \leq y \leq d$, y C_1^+ se puede parametrizar por $y \rightsquigarrow (\psi_1(y), y)$, $c \leq y \leq d$, luego,

$$\int_{C^+} Q dy = \int_{C_1^- + B_1^+ + C_2^+ + B_2^-} Q dy = \int_{C_2^+} Q dy + \int_{C_1^-} Q dy.$$

Aplicando el teorema de Fubini y el teorema fundamental de cálculo:

$$\begin{aligned} \int_D \frac{\partial Q}{\partial x} dx dy &= \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} \frac{\partial Q}{\partial x} dx dy = \int_c^d [Q(\psi_2(y), y) - Q(\psi_1(y), y)] dy = \\ &= \int_{C_2^+} Q dy - \int_{C_1^-} Q dy = \int_{C_2^+} Q dy + \int_{C_1^+} Q dy = \int_{C^+} Q dy. \end{aligned}$$

c.q.d.

La orientación correcta (positiva) para las curvas frontera de una región D se puede recordar mediante esta regla: si caminamos a lo largo de la curva C con la orientación correcta, la región D estará a nuestra izquierda.

El teorema de Green se aplica en realidad a cualquier región "decente" en \mathbb{R}^2 ; esto es, para regiones que no son del tipo 3, pero que se pueden descomponer en partes, cada una del tipo 3.

Usemos la notación ∂D para la curva orientada C^+ , esto es, la curva frontera de D orientada en el sentido correcto. Entonces podemos escribir el teorema de Green como

$$\int_{\partial D} P dx + Q dy = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

El teorema de Green es muy útil, pues relaciona una integral de línea alrededor de la frontera de una región, con una integral de área sobre el interior de la región, y en

13.7. TEOREMA DE GREEN.

muchos casos es más fácil evaluar la integral de línea que la integral de área. Por ejemplo, si sabemos que P se anula en la frontera, podemos concluir de manera inmediata que $\int_D \frac{\partial P}{\partial y} dx dy = 0$ aunque $\frac{\partial P}{\partial y}$ no necesariamente se anule en el interior.

Ejemplo. Verifica el teorema de Green para $P(x, y) = x$ y $Q(x, y) = xy$ donde D es el disco unitario $x^2 + y^2 \leq 1$.

Solución:

Evaluamos directamente ambos lados en el teorema de Green. La frontera de D es el círculo unitario parametrizado por

$$x = \cos t, \quad y = \sin t, \quad \text{con } 0 \leq t \leq 2\pi,$$

de modo que

$$\int_{\partial D} P dx + Q dy = \int_0^{2\pi} [(\cos t)(-\sin t) + \cos t \sin t \cos t] dt = 0.$$

Por otro lado,

$$\int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_D y dx dy$$

que es cero por simetría. Así, se verifica el teorema de Green en este caso.

Consideremos una región Ω en forma de corona circular: su frontera consiste en dos curvas C_1 y C_2 . No podemos aplicar directamente el teorema de Green a Ω pero podemos dividir Ω en dos regiones y aplicar luego el teorema de Green a cada trozo.



Con Ω_1 y Ω_2 como en la figura, tenemos

$$\int_{\Omega_1} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\partial \Omega_1} P dx + Q dy,$$

$$\int_{\Omega_2} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\partial \Omega_2} P dx + Q dy.$$

Cuando sumamos las dos integrales dobles, obtenemos la integral doble sobre Ω . Cuando sumamos las integrales de línea, las integrales sobre los cortes transversales se anulan y nos quedamos con la integral sobre C_1 , en sentido contrario al de las agujas del reloj y con la integral sobre C_2 en el mismo sentido que el de las agujas del reloj. Luego para la región en forma de corona circular

$$\int_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \oint_{C_1^+} P dx + Q dy + \oint_{C_2^-} P dx + Q dy.$$

Como corolario a esto vemos que si $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ sobre Ω la integral doble de la izquierda es nula, y la suma de las integrales de la derecha es cero; por lo tanto

$$\oint_{C_1^+} P dx + Q dy = - \oint_{C_2^-} P dx + Q dy = \oint_{C_2^+} P dx + Q dy. \quad (13.35)$$

Ejemplo. Sea C_1 una curva simple que no pasa por el origen $(0, 0)$. Demuestra que

$$\oint_{C_1} -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = \begin{cases} 0, & \text{si } C_1 \text{ no encierra el origen;} \\ 2\pi, & \text{si } C_1 \text{ encierra el origen.} \end{cases}$$

Solución:

En este caso, como vimos en un ejemplo de la sección 13.1.

$$\frac{\partial P}{\partial y}(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial Q}{\partial x}(x, y), \quad \text{excepto en el origen.}$$

Si el origen no se encuentra en la región encerrada por C_1 , se verifica que

$$\frac{\partial P}{\partial y}(x, y) = \frac{\partial Q}{\partial x}(x, y)$$

sobre dicha región, luego, por el teorema de Green 13.11, la integral de línea es nula.

Si el origen está en la región encerrada por C_1 , trazamos en dicha región una pequeña circunferencia centrada en el origen C_2 : $x^2 + y^2 = a^2$.

Dado que $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0$ en la región con forma de corona circular limitada por C_1 y C_2 , sabemos por la fórmula 13.35 que la integral de línea sobre C_1 es igual a la integral de línea sobre C_2 . Sólo nos queda por demostrar que la integral de línea sobre C_2 vale 2π . Para conseguirlo, parametrizamos la circunferencia según

$$\vec{r}(\theta) = a \cos \theta \vec{i} + a \sin \theta \vec{j}, \quad \text{con } 0 \leq \theta \leq 2\pi.$$

Tenemos entonces

$$\oint_{C_2} -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = \int_0^{2\pi} (\sin^2 \theta + \cos^2 \theta) d\theta = 2\pi.$$

Podemos usar el teorema de Green 13.11 para obtener una fórmula para el área de una región acotada por una curva cerrada simple.

Teorema 13.12 Si C es una curva cerrada simple que acota una región para la cual se aplica el teorema de Green, entonces el área de la región D acotada por $C = \partial D$ es

$$A(D) = \frac{1}{2} \int_{\partial D} x dy - y dx \quad (13.36)$$

13.7. TEOREMA DE GREEN.

Demostración:

Sean $P(x, y) = -y$ y $Q(x, y) = x$; entonces por el teorema de Green tenemos

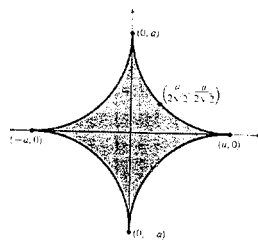
$$\frac{1}{2} \int_{\partial D} x dy - y dx = \frac{1}{2} \int_D \left[\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y} \right] dx dy = \frac{1}{2} \int_D 2 dx dy = \int_D dx dy = A(D).$$

c.q.d.

Ejemplo. Calcula el área de la región encerrada por la hipocicloide definida por la ecuación $x^{2/3} + y^{2/3} = a^{2/3}$ usando la parametrización

$$x = a \cos^3 \theta, \quad y = a \sin^3 \theta,$$

$$\text{con } 0 \leq \theta \leq 2\pi$$



Solución:

$$\begin{aligned} A &= \frac{1}{2} \int_{\partial D} x dy - y dx = \\ &= \frac{1}{2} \int_0^{2\pi} [(a \cos^3 \theta) (3a \sin^2 \theta \cos \theta) - (a \sin^3 \theta) (3a \cos^2 \theta \sin \theta)] d\theta = \\ &= \frac{3}{2} a^2 \int_0^{2\pi} (\sin^2 \theta \cos^4 \theta + \sin^4 \theta \cos^2 \theta) d\theta = \frac{3}{2} a^2 \int_0^{2\pi} \sin^2 \theta \cos^2 \theta d\theta = \\ &= \frac{3}{8} a^2 \int_0^{2\pi} \sin^2 2\theta d\theta = \frac{3}{8} a^2 \int_0^{2\pi} \frac{1 - \cos 4\theta}{2} d\theta = \\ &= \frac{3}{16} a^2 \int_0^{2\pi} d\theta - \frac{3}{16} a^2 \int_0^{2\pi} \cos 4\theta d\theta = \frac{3}{8} a^2 \pi. \end{aligned}$$

La forma del enunciado del teorema de Green contenido en el teorema 13.11 no es la que generalizaremos en las secciones siguientes. Podemos reescribir con elegancia el teorema, en lenguaje de campos vectoriales.

Teorema 13.13 Forma vectorial del Teorema de Green.

Sea $D \subset \mathbb{R}^2$ una región de tipo 3 y sea ∂D su frontera (orientada en sentido contrario al que giran las manecillas del reloj). Sea $\vec{F} = P\vec{i} + Q\vec{j}$ un campo vectorial C^1 en D . Entonces

$$\int_{\partial D} \vec{F} \cdot d\vec{s} = \int_D (\text{rot } \vec{F}) \cdot \vec{k} dA = \int_D (\vec{\nabla} \times \vec{F}) \cdot \vec{k} dA. \quad (13.37)$$

Este resultado se sigue fácilmente del teorema 13.11 después de interpretar los diferentes símbolos.

Ejemplo. Sea $\vec{F} = (xy^2, y+x)$. Integra $(\vec{\nabla} \times \vec{F}) \cdot \vec{k}$ sobre la región del primer cuadrante acotado por las curvas $y = x^2$ e $y = x$.

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

Solución:

Método 1. Aquí calculamos

$$\vec{\nabla} \times \vec{F} = \left(0, 0, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) = (1 - 2xy) \vec{k}.$$

Así, $(\vec{\nabla} \times \vec{F}) \cdot \vec{k} = 1 - 2xy$. Esto se puede integrar sobre la región dada D usando una integral iterada como sigue:

$$\iint_D (\vec{\nabla} \times \vec{F}) \cdot \vec{k} \, dx \, dy = \int_0^1 \int_{x^2}^x (1 - 2xy) \, dy \, dx = \frac{1}{12}.$$

Método 2. Aquí usamos el teorema 13.13 para obtener

$$\iint_D (\vec{\nabla} \times \vec{F}) \cdot \vec{k} \, dx \, dy = \int_{\partial D} \vec{F} \cdot d\vec{s}.$$

La integral de línea de \vec{F} a lo largo de la curva $y = x$ de izquierda a derecha es

$$\int_0^1 F_1 \, dx + F_2 \, dy = \int_0^1 (x^3 + 2x) \, dx = \frac{5}{4}.$$

A lo largo de la curva $y = x^2$ obtenemos

$$\int_0^1 F_1 \, dx + F_2 \, dy = \int_0^1 x^5 \, dx + (x + x^2)(2x \, dx) = \frac{4}{3}.$$

Y, recordando que la integral a lo largo de $y = x$ se toma de derecha a izquierda,

$$\int_{\partial D} \vec{F} \cdot d\vec{s} = \frac{4}{3} - \frac{5}{4} = \frac{1}{12}.$$

Ejercicios.

1) Verifica el teorema de Green para el disco D con centro en $(0,0)$ y radio R y las funciones

a) $P(x, y) = x + y, \quad Q(x, y) = y;$

b) $P(x, y) = xy, \quad Q(x, y) = xy;$

c) $P(x, y) = -y^3, \quad Q(x, y) = x^3.$

2) Evalúa $\oint_C (2x^3 - y^3) \, dx + (x^3 + y^3) \, dy$, donde C es el círculo unitario, y verifica en este caso el teorema de Green.

13.8. TEOREMA DE STOKES.

- 3) Verifica el teorema de Green para la integral curvilínea $\oint_C 2(x^2 + y^2) dx + (x + y)^2 dy$, donde C es el contorno del triángulo de vértices $A(1, 1)$, $B(2, 2)$ y $C(1, 3)$, recorrido en sentido positivo.
- 4) Usa el teorema de Green para calcular $\oint_C xy^2 dx + x^2 y dy$, donde C es la circunferencia $x^2 + y^2 = a^2$.
- 5) Usa el teorema de Green para calcular $\oint_C (1 + 10xy + y^2) dx + (6xy + 5x^2) dy$, donde C es el cuadrado de vértices $(0, 0)$, $(a, 0)$, (a, a) , $(0, a)$.
- 6) Usa el teorema de Green para calcular $\oint_C (4xy + y^2) dx + (xy + 3x^2) dy$, donde C es la circunferencia $(x + 1)^2 + (y - 2)^2 = 9$.
- 7) Prueba las identidades de Green

$$\int_{\partial\Omega} f \bar{\nabla} g \cdot \bar{n} dS = \int_{\Omega} [f \nabla^2 g + \bar{\nabla} f \cdot \bar{\nabla} g] dV,$$

$$\int_{\partial\Omega} [f \bar{\nabla} g - g \bar{\nabla} f] \cdot \bar{n} dS = \int_{\Omega} [f \nabla^2 g - g \nabla^2 f] dV.$$

13.8 Teorema de Stokes.

El teorema de Stokes relaciona la integral de línea de un campo vectorial alrededor de una curva cerrada simple C en \mathbb{R}^3 , con la integral sobre una superficie S de la cual C es la frontera. En este aspecto se parece mucho al teorema de Green.

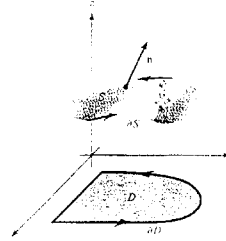
Sea S la superficie descrita por $z = f(x, y) = f(u, v)$ con $(u, v) \in D$. Podemos parametrizar S por $\bar{\Phi}: D \rightarrow \mathbb{R}^3$ dada por $\bar{\Phi}(x, y) = (x, y, f(x, y))$. En este caso, vimos en la sección 13.6 que la integral $\int_S \bar{F} \cdot d\bar{S}$ se puede escribir de manera más simple, dada por la fórmula (13.33)

$$\int_S \bar{F} \cdot d\bar{S} = \int_D \left[F_1 \left(-\frac{\partial f}{\partial x} \right) + F_2 \left(-\frac{\partial f}{\partial y} \right) + F_3 \right] dx dy, \quad (13.33)$$

donde $\bar{F} = (F_1, F_2, F_3)$. En la sección 13.7 supusimos que las regiones D consideradas eran del tipo 3 para poder dar la demostración del teorema de Green, pero notamos que el teorema es válido para una clase más amplia de regiones. En esta sección supondremos que D es una región cuya frontera es una curva cerrada simple a la cual se puede aplicar el teorema de Green. Según se explicó en la sección 13.7, para aplicar el teorema de Green se necesita escoger una orientación de la frontera de D : la orientación que haga que se cumpla el teorema se llamará *positiva*.

Recordad que si D es del tipo 3, entonces la orientación positiva es en sentido contrario al que giran las manecillas del reloj. Supongamos que $\vec{\sigma}: [a, b] \rightarrow \mathbb{R}^2$, con $\vec{\sigma}(t) = (x(t), y(t))$ sea una parametrización de ∂D en dirección positiva. Definimos entonces la curva frontera ∂S como la curva cerrada simple orientada que es la imagen de la función $\vec{\eta}: t \rightsquigarrow (x(t), y(t), f(x(t), y(t)))$ con la orientación inducida por $\vec{\eta}$.

Para recordar la dirección positiva de ∂S , imaginemos un “observador” caminando a lo largo de la frontera de la superficie donde la normal apunta para el mismo lado que su cabeza; se estará moviendo en la dirección positiva si la superficie está a su izquierda: esta orientación de ∂S suele llamarse orientación inducida por una normal \vec{n} “hacia arriba”.



Ahora estamos preparados para enunciar y probar uno de los resultados fundamentales de esta sección.

Teorema 13.14 Teorema de Stokes para gráficas.

Sea S la superficie orientada definida por una función C^2 , $z = f(x, y)$, con $(x, y) \in D$. y sea \vec{F} un campo vectorial C^1 en S . Entonces, si ∂S denota la curva frontera orientada de S según se ha definido arriba, se cumplirá que

$$\int_S \text{rot } \vec{F} \cdot d\vec{S} = \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s}. \quad (13.38)$$

Debemos recordar que $\int_{\partial S} \vec{F} \cdot d\vec{s}$ es la integral alrededor de ∂S de la componente tangencial de \vec{F} , mientras que $\int_S \vec{G} \cdot d\vec{S}$ es la integral sobre S de $\vec{G} \cdot \vec{n}$, la componente normal de \vec{G} . Así, el teorema de Stokes dice que **la integral de la componente normal del rotacional de un campo vectorial \vec{F} sobre una superficie S , es igual a la integral de la componente tangencial de \vec{F} alrededor de la frontera ∂S .**

Demostración:

Si $\vec{F} = F_1 \vec{i} + F_2 \vec{j} + F_3 \vec{k}$, entonces

$$\text{rot } \vec{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \vec{i} - \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) \vec{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \vec{k}.$$

Por lo tanto, usamos la fórmula (13.33) para escribir

$$\int_S \text{rot } \vec{F} \cdot d\vec{S} = \int_D \left[-\frac{\partial f}{\partial x} \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) + \frac{\partial f}{\partial y} \left(\frac{\partial F_3}{\partial x} - \frac{\partial F_1}{\partial z} \right) + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \right] dA$$

Por otro lado,

$$\int_{\partial S} \vec{F} \cdot d\vec{s} = \int_{\vec{\eta}} \vec{F} \cdot d\vec{s} = \int_{\vec{\eta}} F_1 dx + F_2 dy + F_3 dz.$$

13.8. TEOREMA DE STOKES.

donde $\vec{\eta}: t \rightsquigarrow (x(t), y(t), f(x(t), y(t)))$ es la parametrización que preserva la orientación de la curva cerrada simple orientada ∂S estudiada anteriormente, siendo $z(t) = f(x(t), y(t))$. Así,

$$\int_{\partial S} \vec{F} \cdot d\vec{s} = \int_a^b \left(F_1 \frac{dx}{dt} + F_2 \frac{dy}{dt} + F_3 \frac{dz}{dt} \right) dt,$$

Pero, por la regla de la cadena

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

Sustituyendo esta expresión en la última ecuación, obtenemos

$$\begin{aligned} \int_{\partial S} \vec{F} \cdot d\vec{s} &= \int_a^b \left[\left(F_1 + F_3 \frac{\partial z}{\partial x} \right) \frac{dx}{dt} + \left(F_2 + F_3 \frac{\partial z}{\partial y} \right) \frac{dy}{dt} \right] dt = \\ &= \int_{\vec{\sigma}} \left(F_1 + F_3 \frac{\partial z}{\partial x} \right) dx + \left(F_2 + F_3 \frac{\partial z}{\partial y} \right) dy = \\ &= \int_{\partial D} \left(F_1 + F_3 \frac{\partial z}{\partial x} \right) dx + \left(F_2 + F_3 \frac{\partial z}{\partial y} \right) dy. \end{aligned}$$

Aplicando el teorema de Green a esta ecuación se obtiene

$$\int_D \left[\frac{\partial}{\partial x} \left(F_2 + F_3 \frac{\partial z}{\partial y} \right) - \frac{\partial}{\partial y} \left(F_1 + F_3 \frac{\partial z}{\partial x} \right) \right] dA.$$

Usamos ahora la regla de la cadena, recordando que F_1 , F_2 y F_3 son funciones de x , y y z , y que z es función de x e y , para obtener

$$\begin{aligned} &\int_D \left(\frac{\partial F_2}{\partial x} + \frac{\partial F_2}{\partial z} \frac{\partial z}{\partial x} + \frac{\partial F_3}{\partial x} \frac{\partial z}{\partial y} + \frac{\partial F_3}{\partial z} \frac{\partial z}{\partial x} \frac{\partial z}{\partial y} + F_3 \frac{\partial^2 z}{\partial x \partial y} \right) dA \\ &- \int_D \left(\frac{\partial F_1}{\partial y} + \frac{\partial F_1}{\partial z} \frac{\partial z}{\partial y} + \frac{\partial F_3}{\partial y} \frac{\partial z}{\partial x} + \frac{\partial F_3}{\partial z} \frac{\partial z}{\partial y} \frac{\partial z}{\partial x} + F_3 \frac{\partial^2 z}{\partial y \partial x} \right) dA. \end{aligned}$$

Los últimos dos términos en cada paréntesis se cancelan entre sí, y podemos reajustar los términos para obtener la expresión que obtuvimos para la integral $\int_S \text{rot } \vec{F} \cdot d\vec{S}$, demostrando así la igualdad

$$\int_S \text{rot } \vec{F} \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s}. \quad (13.38)$$

c.q.d.

Ejemplo 1. Comprueba el teorema de Stokes para $\vec{F} = -3y\vec{i} + 3x\vec{j} + z^4\vec{k}$ tomando como S la parte del elipsoide $2x^2 + 2y^2 + z^2 = 1$ que está por encima del plano $z = \frac{1}{\sqrt{2}}$.

Solución:

Claramente S es la gráfica de $z = f(x, y) = \sqrt{1 - 2(x^2 + y^2)}$ con (x, y) restringidos al disco $D: x^2 + y^2 \leq \frac{1}{4}$. Ahora

$$\vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -3y & 3x & z^4 \end{vmatrix} = \left[\frac{\partial}{\partial x}(3x) + \frac{\partial}{\partial y}(3y) \right] \vec{k} = 6\vec{k}.$$

Usando la fórmula (13.33)

$$\int_S \vec{F} \cdot d\vec{S} = \int_D \left[F_1 \left(-\frac{\partial f}{\partial x} \right) + F_2 \left(-\frac{\partial f}{\partial y} \right) + F_3 \right] dx dy, \quad (13.33)$$

obtenemos

$$\int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} = \int_D 6 dx dy = 6(\text{área de } D) = \frac{3}{2}\pi,$$

siendo D el círculo $2x^2 + 2y^2 = \frac{1}{2}$, esto es el círculo con centro en $(0, 0)$ y radio $\frac{1}{2}$, cuya área es $\pi r^2 = \frac{1}{4}\pi$. Vamos ahora a comprobar el segundo miembro de la igualdad (13.38).

La curva C que limita es el conjunto de todos los (x, y, z) tales que $x^2 + y^2 \leq \frac{1}{4}$. Podemos parametrizar C haciendo

$$\vec{\sigma}(\theta) = \frac{1}{2} \cos \theta \vec{i} + \frac{1}{2} \sin \theta \vec{j} + \frac{1}{\sqrt{2}} \vec{k}, \quad \text{con } \theta \in [0, 2\pi].$$

Usando la definición de integral de línea, fórmula (13.10)

$$\int_{\vec{\sigma}} \vec{F} \cdot d\vec{s} = \int_a^b \vec{F}(\vec{\sigma}(t)) \cdot \vec{\sigma}'(t) dt.$$

obtenemos

$$\begin{aligned} \int_{\partial S} \vec{F} \cdot d\vec{s} &= \int_0^{2\pi} \left(-\frac{3}{2} \sin \theta \vec{i} + \frac{3}{2} \cos \theta \vec{j} + z^4 \vec{k} \right) \cdot \left(-\frac{1}{2} \sin \theta \vec{i} + \frac{1}{2} \cos \theta \vec{j} \right) d\theta = \\ &= \int_0^{2\pi} \left(\frac{3}{4} \sin^2 \theta + \frac{3}{4} \cos^2 \theta \right) d\theta = \frac{3}{4} \int_0^{2\pi} d\theta = \frac{3}{2}\pi. \end{aligned}$$

Este es el valor obtenido para la integral de superficie.

Ejemplo 2. Comprueba el teorema de Stokes para $\vec{F} = z^2\vec{i} - 2x\vec{j} + y^3\vec{k}$ tomando para S la mitad superior de la esfera unidad $x^2 + y^2 + z^2 = 1$.

13.8. TEOREMA DE STOKES.

Solución:

$$\vec{\nabla} \times \vec{F} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ z^2 & -2x & y^3 \end{vmatrix} = 3y^2 \vec{i} + 2z \vec{j} - 2\vec{k}.$$

Usaremos ahora el resultado del teorema 13.10, esto es la fórmula (13.32)

$$\int_S \vec{F} \cdot d\vec{S} = \int_S (\vec{F} \cdot \vec{n}) \, dS. \quad (13.32)$$

Usando la normal unitaria superior dada por

$$\vec{n} = x \vec{i} + y \vec{j} + z \vec{k},$$

obtenemos

$$\begin{aligned} \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} &= \int_S [(\vec{\nabla} \times \vec{F}) \cdot \vec{n}] \, dS = \\ &= \int_S [(3y^2 \vec{i} + 2z \vec{j} - 2\vec{k}) \cdot (x \vec{i} + y \vec{j} + z \vec{k})] \, dS = \\ &= \int_S (3xy^2 + 2yz - 2z) \, dS = -2\pi, \end{aligned}$$

donde la primera integral se anula por ser S simétrica con respecto al plano yz y el integrando impar con respecto a x : la segunda integral se anula por ser S simétrica con respecto al plano xz y el integrando impar con respecto a y .

Ejemplo 3. Usa el teorema de Stokes para evaluar la integral de línea

$$\int_C -y^3 \, dx + x^3 \, dy - z^3 \, dz,$$

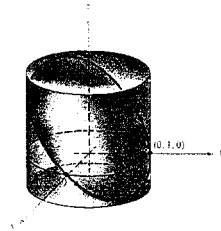
donde C es la intersección del cilindro $x^2 + y^2 = 1$ y el plano $x + y + z = 1$, y la orientación de C corresponde al movimiento en sentido contrario al que giran las manecillas del reloj, en el plano xy .

Solución:

La curva C acota la superficie S definida por $z = 1 - x - y = f(x, y)$ para (x, y) en $D = \{(x, y) / x^2 + y^2 \leq 1\}$. Ponemos $\vec{F} = -y^3 \vec{i} + x^3 \vec{j} + z^3 \vec{k}$ que tiene rotacional $\vec{\nabla} \times \vec{F} = (3x^2 + 3y^2) \vec{k}$.

Entonces, por el teorema de Stokes, la integral de línea es igual a la integral de superficie

$$\int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S}.$$



Pero $\vec{\nabla} \times \vec{F}$ tiene sólo componente \vec{k} . Así, por la fórmula (13.33), tenemos

$$\int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} = \int_D (3x^2 + 3y^2) dx dy.$$

Esta integral se puede evaluar pasando a coordenadas polares. Al hacerlo, obtenemos

$$3 \int_D (x^2 + y^2) dx dy = 3 \int_0^1 \int_0^{2\pi} r^2 \cdot r d\theta dr = 6\pi \int_0^1 r^3 dr = \frac{3}{2}\pi.$$

Verifiquemos este resultado evaluando directamente la integral de línea

$$\int_C -y^3 dx + x^3 dy - z^3 dz.$$

Podemos parametrizar la curva ∂D por las ecuaciones

$$x = \cos \theta, \quad y = \text{sen } \theta, \quad z = 0, \quad \text{con } 0 \leq \theta \leq 2\pi.$$

Entonces la curva C está parametrizada por las ecuaciones

$$x = \cos \theta, \quad y = \text{sen } \theta, \quad z = 1 - \text{sen } \theta - \cos \theta, \quad \text{con } 0 \leq \theta \leq 2\pi.$$

Así, $\int_C -y^3 dx + x^3 dy - z^3 dz$ se transforma en

$$\int_0^{2\pi} [(-\text{sen}^3 \theta)(-\text{sen } \theta) + \cos^3 \theta \cos \theta - (1 - \text{sen } \theta - \cos \theta)^3 (\text{sen } \theta - \cos \theta)] dt$$

que con unos cálculos un poco laboriosos nos da el valor previamente obtenido.

Para simplificar la demostración del anterior teorema de Stokes, supusimos que la superficie S podría describirse como la gráfica de una función $z = f(x, y)$, con $(x, y) \in D$, donde D es alguna región a la que se aplica el teorema de Green.

Sin embargo, sin mucho más esfuerzo podemos obtener un teorema más general para superficies parametrizadas orientadas S . La dificultad principal radica en la definición de ∂S . Suponer que $\vec{\Phi}: D \rightarrow \mathbb{R}^3$ es una parametrización de una superficie S y $\vec{\sigma}(t) = (u(t), v(t))$ es una parametrización de ∂D . Podríamos sentirnos tentados a definir ∂S como la curva parametrizada por $t \rightsquigarrow \vec{\eta}(t) = \vec{\Phi}(u(t), v(t))$. Sin embargo, con esta definición, ∂S podría no ser la frontera de S en ningún sentido geométrico razonable.

Por ejemplo, llegaríamos a la conclusión de que la frontera de la esfera unitaria S parametrizada mediante coordenadas esféricas en \mathbb{R}^3 , es la mitad del gran círculo en S que está en el plano xz , pero está claro que en un sentido geométrico S es una superficie suave (ni puntas ni cúspides) sin fronteras ni lados. Así, este gran círculo es, en cierto sentido, la frontera "falsa" de S .

Podemos eludir esta dificultad suponiendo que $\vec{\Phi}$ es uno a uno en todo D . Entonces la imagen de ∂D bajo $\vec{\Phi}$, a saber, $\vec{\Phi}(\partial D)$, será la frontera geométrica de $S = \vec{\Phi}(D)$. Si $\vec{\sigma}(t) = (u(t), v(t))$ es una parametrización de ∂D en dirección positiva, definiremos ∂S como la curva cerrada simple orientada que es la imagen de la función $\vec{\eta}: t \rightsquigarrow \vec{\Phi}(u(t), v(t))$ con la orientación de ∂S inducida por $\vec{\eta}$.

13.8. TEOREMA DE STOKES.

Teorema 13.15 Teorema de Stokes para superficies parametrizadas.

Sea S una superficie orientada definida por una parametrización $\vec{\Phi}: D \rightarrow S$, uno a uno. Denotemos por ∂S la frontera orientada de S y sea \vec{F} un campo vectorial C^1 en S . Entonces valdrá la ecuación (13.38)

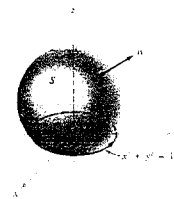
$$\int_S \text{rot } \vec{F} \cdot d\vec{S} = \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s}. \quad (13.38)$$

Esto se demuestra de la misma manera que el teorema 13.14.

Ejemplo. Sea S la superficie mostrada en la figura, con la orientación indicada. Sea $\vec{F} = y\vec{i} - x\vec{j} + e^{xy}\vec{k}$. Evalúa $\int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S}$.

Solución:

Ésta es una superficie parametrizada y pudo ser parametrizada usando coordenadas esféricas. Sin embargo, no necesitamos hallar explícitamente $\vec{\Phi}$ para resolver este problema. Por el teorema 13.15,



$$\int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s},$$

de modo que si parametrizamos ∂S por

$$x(\theta) = \cos \theta, \quad y(\theta) = \sin \theta, \quad \text{con } 0 \leq \theta \leq 2\pi,$$

determinamos

$$\begin{aligned} \int_{\partial S} \vec{F} \cdot d\vec{s} &= \int_0^{2\pi} \left(y \frac{dx}{d\theta} - x \frac{dy}{d\theta} \right) d\theta = \int_0^{2\pi} (-\sin^2 \theta - \cos^2 \theta) d\theta = \\ &= - \int_0^{2\pi} d\theta = -2\pi = \int_S (\vec{\nabla} \times \vec{F}) \cdot d\vec{S}. \end{aligned}$$

Usemos ahora el teorema de Stokes para justificar la interpretación física de $\vec{\nabla} \times \vec{F}$ en términos de ruedas con aspas propuesta en la sección 13.3. Recordando el teorema 13.15, tenemos

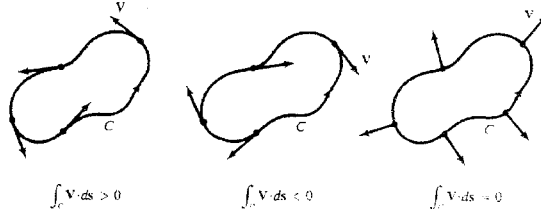
$$\int_S (\text{rot } \vec{F}) \cdot \vec{n} dS = \int_S (\text{rot } \vec{F}) \cdot d\vec{S} = \int_{\partial S} \vec{F} \cdot d\vec{s} = \int_{\partial S} F_T ds,$$

donde F_T es la componente tangencial de \vec{F} . Esto significa que la integral de la componente normal del rotacional de un campo vectorial sobre una superficie orientada S , es igual a la integral de línea de \vec{F} a lo largo de ∂S , lo cual, a su vez, es igual a la integral de trayectoria de la componente tangencial de \vec{F} sobre ∂S .

Supongamos que \vec{V} representa el campo vectorial de velocidad de un fluido. Consideremos un punto P y un vector unitario \vec{n} . Denotemos por S_ρ el disco de radio ρ y centro P , el cual es perpendicular a \vec{n} . Por el teorema de Stokes,

$$\int_{S_\rho} \text{rot } \vec{V} \cdot d\vec{S} = \int_{S_\rho} \text{rot } \vec{V} \cdot \vec{n} dS = \int_{\partial S_\rho} \vec{V} \cdot d\vec{s},$$

donde ∂S_p tiene la orientación inducida por \vec{n} . Supongamos, por ejemplo, que \vec{V} apunta en dirección tangente a la curva orientada C .



Entonces, claramente $\int_C \vec{V} \cdot d\vec{s} > 0$, y las partículas en C tienden a rotar en sentido contrario al que giran las manecillas del reloj. Si \vec{V} apunta en dirección opuesta, $\int_C \vec{V} \cdot d\vec{s} < 0$. Si \vec{V} es perpendicular a C , entonces las partículas no giran en C y $\int_C \vec{V} \cdot d\vec{s} = 0$.

En general, al ser $\int_C \vec{V} \cdot d\vec{s}$ la integral de la componente tangencial de \vec{V} , representa la cantidad neta de giro del fluido en dirección contraria a la que giran las manecillas del reloj alrededor de C . Por lo tanto nos referimos a $\int_C \vec{V} \cdot d\vec{s}$ como la **circulación** de \vec{V} alrededor de C , y el vector $\text{rot } \vec{V}$ es el **vector vorticidad**.

Ejemplo: Ley de Faraday

Una ley básica de la teoría electromagnética es que si $\vec{E}(t, x, y, z)$ y $\vec{B}(t, v, y, z)$ representan los campos eléctrico y magnético, respectivamente, en el tiempo t , entonces

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

donde $\vec{\nabla} \times \vec{E}$ se calcula manteniendo t fija y $\frac{\partial \vec{B}}{\partial t}$ se calcula manteniendo x, y y z constantes. Usemos el teorema de Stokes para determinar lo que esta ley significa físicamente. Supongamos que S es una superficie a la que se aplica el teorema de Stokes. Entonces

$$\oint_{\partial S} \vec{E} \cdot d\vec{s} = \int_S \text{rot } \vec{E} \cdot d\vec{S} = - \int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} = -\frac{\partial}{\partial t} \int_S \vec{B} \cdot d\vec{S}.$$

(La última igualdad se puede justificar si \vec{B} es de clase C^1 .) Así, obtenemos

$$\oint_{\partial S} \vec{E} \cdot d\vec{s} = -\frac{\partial}{\partial t} \int_S \vec{B} \cdot d\vec{S}.$$

Esta igualdad se conoce como **ley de Faraday**. La cantidad $\oint_{\partial S} \vec{E} \cdot d\vec{s}$ representa la fuerza electromotriz (voltaje) en ∂S , y si ∂S fuera un alambre, una corriente eléctrica fluiría en

13.9. TEOREMA DE GAUSS.

proporción a este voltaje. Además, $\int_S \vec{B} \cdot d\vec{S}$ se llama flujo de \vec{B} , o flujo magnético. Así, la ley de Faraday dice que el voltaje inducido en un circuito cerrado es igual a la tasa de cambio del flujo magnético a través del circuito, y se opone al cambio del mismo flujo (de ahí el signo $-$, hecho que se conoce como ley de Lenz).

Ejercicios.

- 1) Aplicando la fórmula de Stokes, halla las integrales que se dan a continuación y comprueba los resultados calculándolas directamente.

a) $\oint_C (y+z) dx + (z+x) dy + (x+y) dz$, donde c es la circunferencia
 $x^2 + y^2 + z^2 = a^2, \quad x + y + z = 0.$

b) $\oint_C (y-z) dx + (z-x) dy + (x-y) dz$, donde c es la elipse
 $x^2 + y^2 = 1, \quad x + z = 1.$

c) $\oint_C y^2 dx + z^2 dy + x^2 dz$, donde c es el contorno del triángulo de vértices
 $A(a, 0, 0), \quad B(0, a, 0), \quad C(0, 0, a).$

- 2) Sea S la mitad superior de la esfera unidad $x^2 + y^2 + z^2 = 1$ y sea \vec{n} la normal unitaria superior. Halla

$$\int_S [(\vec{\nabla} \times \vec{v}) \cdot \vec{n}] dS$$

- (a) mediante un cálculo directo; (b) aplicando el teorema de Stokes;

para los campos

- i) $\vec{v}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k};$ ii) $\vec{v}(x, y, z) = y\vec{i} - x\vec{j} + z\vec{k};$
 iii) $\vec{v}(x, y, z) = z^2\vec{i} + 2x\vec{j} - y^3\vec{k};$ iv) $\vec{v}(x, y, z) = 6xz\vec{i} - x^2\vec{j} - 3y^2\vec{k}.$

13.9 Teorema de Gauss.

El teorema de Gauss asegura que el flujo de un campo vectorial hacia afuera de una superficie cerrada es igual a la integral de la divergencia de ese campo vectorial sobre el volumen encerrado por la superficie. Se trata de un resultado paralelo al teorema de Stokes y al de Green, en el sentido de que relaciona una integral sobre un objeto geométrico cerrado (curva o superficie) con una integral sobre una región contenida (superficie o volumen).

La frontera de una región tipo I, II o III en \mathbb{R}^3 es una superficie formada por un número finito (a lo más seis, por lo menos dos) de superficies que se pueden describir como gráficas de funciones de \mathbb{R}^2 a \mathbb{R} . Este tipo de superficie se llama superficie cerrada. Las superficies que componen dicha superficie cerrada se llaman "caras". Por ejemplo, el cubo es una

región del tipo IV, esto es, es simultáneamente del tipo I, II y III; la esfera es la frontera de una bola sólida, que es también una región del tipo IV.

Las superficies cerradas se pueden orientar de dos maneras. En la primera, orientación exterior, la normal apunta hacia afuera en el espacio. y en la segunda, la orientación interior, la normal apunta hacia adentro de la región.

Supongamos que S sea una superficie cerrada orientada de alguna de estas dos maneras y \vec{F} un campo vectorial en S . Entonces, como definimos en la sección 13.6

$$\int_S \vec{F} \cdot d\vec{S} = \sum_i \int_{S_i} \vec{F} \cdot d\vec{S}.$$

Si S tiene la orientación exterior, la integral $\int_S \vec{F} \cdot d\vec{S}$ mide el flujo total de \vec{F} hacia afuera a través de S . Esto es, si pensamos \vec{F} como el campo de velocidad de un fluido, $\int_S \vec{F} \cdot d\vec{S}$ indica la cantidad de fluido que sale de la región S por unidad de tiempo. Si S tiene la orientación interior, la integral $\int_S \vec{F} \cdot d\vec{S}$ mide el flujo total de \vec{F} hacia adentro a través de S .

Recordemos otra forma común de escribir estas integrales de superficie, forma que especifica explícitamente la orientación de S . Sea la orientación de S dada por un vector normal unitario $\vec{n}(x, y, z)$ en cada punto de S . Entonces, tenemos la integral orientada

$$\int_S \vec{F} \cdot d\vec{S} = \int_S (\vec{F} \cdot \vec{n}) dS,$$

esto es, la integral de la componente normal de \vec{F} sobre S . En el resto de la sección, si S es una superficie cerrada que engloba una región Ω , adoptemos el convenio de que $S = \partial\Omega$ tiene dada la orientación exterior, con normal unitaria exterior $\vec{n}(x, y, z)$ en cada punto $(x, y, z) \in S$. Más aún, denotemos la superficie con la orientación opuesta (interior) por $\partial\Omega_{op}$. Entonces la dirección normal unitaria asociada a esta orientación es $-\vec{n}$. Así.

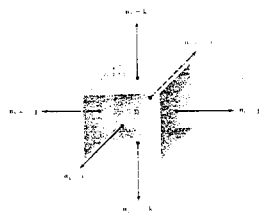
$$\int_{\partial\Omega} \vec{F} \cdot d\vec{S} = \int_S (\vec{F} \cdot \vec{n}) dS = - \int_S (\vec{F} \cdot (-\vec{n})) dS = - \int_{\partial\Omega_{op}} \vec{F} \cdot d\vec{S}.$$

Ejemplo. El cubo unitario Ω dado por

$$0 \leq x \leq 1, \quad 0 \leq y \leq 1, \quad 0 \leq z \leq 1.$$

es una región en el espacio, del tipo IV. Escribimos las caras como

$$\begin{aligned} S_1 : & z = 0, & 0 \leq x \leq 1, & 0 \leq y \leq 1. \\ S_2 : & z = 1, & 0 \leq x \leq 1, & 0 \leq y \leq 1. \\ S_3 : & x = 0, & 0 \leq y \leq 1, & 0 \leq z \leq 1. \\ S_4 : & x = 1, & 0 \leq y \leq 1, & 0 \leq z \leq 1. \\ S_5 : & y = 0, & 0 \leq x \leq 1, & 0 \leq z \leq 1. \\ S_6 : & y = 1, & 0 \leq x \leq 1, & 0 \leq z \leq 1. \end{aligned}$$



13.9. TEOREMA DE GAUSS.

En la figura vemos que

$$\vec{n}_2 = \vec{k} = -\vec{n}_1, \quad \vec{n}_4 = \vec{i} = -\vec{n}_3, \quad \vec{n}_6 = \vec{j} = -\vec{n}_5,$$

de modo que para un campo vectorial continuo $\vec{F} = F_1 \vec{i} + F_2 \vec{j} + F_3 \vec{k}$,

$$\begin{aligned} \int_{\partial\Omega} \vec{F} \cdot d\vec{S} &= \int_S (\vec{F} \cdot \vec{n}) dS = \\ &= - \int_{S_1} F_3 dS + \int_{S_2} F_3 dS - \int_{S_3} F_1 dS + \int_{S_4} F_1 dS - \int_{S_5} F_2 dS + \int_{S_6} F_2 dS. \end{aligned}$$

Llegamos ahora al último de los tres teoremas centrales de este capítulo. Este teorema relaciona integrales de superficie con integrales de volumen; en otras palabras, el teorema asegura que si Ω es una región en \mathbb{R}^3 , entonces el flujo de un campo \vec{F} hacia el exterior a través de la superficie cerrada $\partial\Omega$ es igual a la integral de $\text{div } \vec{F}$ sobre Ω .

Teorema 13.16 Teorema de la divergencia de Gauss.

Sea Ω una región en el espacio, del tipo IV. Denotemos por $\partial\Omega$ la superficie cerrada orientada que acota a Ω . Sea \vec{F} un campo vectorial suave definido en Ω . Entonces

$$\int_{\Omega} (\vec{\nabla} \cdot \vec{F}) dV = \int_{\Omega} (\text{div } \vec{F}) dV = \int_{\partial\Omega} \vec{F} \cdot d\vec{S} = \int_{\partial\Omega} (\vec{F} \cdot \vec{n}) dS. \quad (13.39)$$

Demostración:

Si $\vec{F} = P \vec{i} + Q \vec{j} + R \vec{k}$, entonces por definición, $\text{div } \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$, de modo que podemos escribir

$$\int_{\Omega} \text{div } \vec{F} dV = \int_{\Omega} \frac{\partial P}{\partial x} dV + \int_{\Omega} \frac{\partial Q}{\partial y} dV + \int_{\Omega} \frac{\partial R}{\partial z} dV$$

Por otro lado, la integral de superficie en cuestión es

$$\begin{aligned} \int_{\partial\Omega} \vec{F} \cdot d\vec{S} &= \int_{\partial\Omega} (\vec{F} \cdot \vec{n}) dS = \int_{\partial\Omega} (P \vec{i} + Q \vec{j} + R \vec{k}) \cdot \vec{n} dS = \\ &= \int_{\partial\Omega} P \vec{i} \cdot \vec{n} dS + \int_{\partial\Omega} Q \vec{j} \cdot \vec{n} dS + \int_{\partial\Omega} R \vec{k} \cdot \vec{n} dS. \end{aligned}$$

El teorema se demuestra si logramos probar las tres igualdades

$$\int_{\partial\Omega} P \vec{i} \cdot \vec{n} dS = \int_{\Omega} \frac{\partial P}{\partial x} dV, \quad \int_{\partial\Omega} Q \vec{j} \cdot \vec{n} dS = \int_{\Omega} \frac{\partial Q}{\partial y} dV, \quad \int_{\partial\Omega} R \vec{k} \cdot \vec{n} dS = \int_{\Omega} \frac{\partial R}{\partial z} dV.$$

Probaremos la última igualdad; las otras dos se pueden probar de manera análoga.

Como Ω es una región del tipo I (así como también de los tipos II y III), existen un par de funciones

$$z = f_1(x, y), \quad z = f_2(x, y),$$

CAPÍTULO 13. INTEGRALES SOBRE TRAYECTORIAS Y SUPERFICIES.

cuyo dominio común es una región elemental D en el plano xy , tal que Ω es el conjunto de todos los puntos (x, y, z) que satisfacen

$$f_1(x, y) \leq z \leq f_2(x, y), \quad \text{con } (x, y) \in D.$$

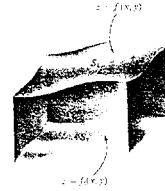
Por la fórmula (12.9) de la sección 12.5, sobre integrales triples, tenemos

$$\int_{\Omega} \frac{\partial R}{\partial z} dV = \int_D \left(\int_{z=f_1(x,y)}^{z=f_2(x,y)} \frac{\partial R}{\partial z} dz \right) dx dy$$

de modo que

$$\int_{\Omega} \frac{\partial R}{\partial z} dV = \int_D [R(x, y, f_2(x, y)) - R(x, y, f_1(x, y))] dx dy.$$

La frontera de Ω es una superficie cerrada cuya tapa S_2 es la gráfica de $z = f_2(x, y)$, $(x, y) \in D$ y cuya parte inferior S_1 es la gráfica de $z = f_1(x, y)$, $(x, y) \in D$. Los otros cuatro lados de $\partial\Omega$ están formados por las superficies S_3, S_4, S_5 y S_6 , cuyas normales son siempre perpendiculares al eje z .



Nótese que pueden faltar algunos de los otros cuatro lados -por ejemplo, si Ω es un bola sólida y $\partial\Omega$ es una esfera- pero esto no afectará el argumento.

Por definición,

$$\int_{\partial\Omega} R \vec{k} \cdot \vec{n} dS = \int_{S_1} R \vec{k} \cdot \vec{n}_1 dS + \int_{S_2} R \vec{k} \cdot \vec{n}_2 dS + \sum_{i=3}^6 \int_{S_i} R \vec{k} \cdot \vec{n}_i dS$$

Como en cada una de las S_3, S_4, S_5 y S_6 la normal \vec{n}_i es perpendicular a \vec{k} , tenemos $\vec{k} \cdot \vec{n}_i = 0$ sobre estas caras, de modo que la integral se reduce a

$$\int_{\partial\Omega} R \vec{k} \cdot \vec{n} dS = \int_{S_1} R \vec{k} \cdot \vec{n}_1 dS + \int_{S_2} R \vec{k} \cdot \vec{n}_2 dS.$$

La superficie S_1 está definida por $z = f_1(x, y)$, de modo que

$$\vec{n}_1 = \frac{\frac{\partial f_1}{\partial x} \vec{i} + \frac{\partial f_1}{\partial y} \vec{j} - \vec{k}}{\sqrt{\left[\frac{\partial f_1}{\partial x}\right]^2 + \left[\frac{\partial f_1}{\partial y}\right]^2 + 1}}$$

(como S_1 es la parte inferior de Ω , para que \vec{n}_1 apunte hacia afuera debe tener componente \vec{k} negativa). Así,

$$\vec{n}_1 \cdot \vec{k} = \frac{-1}{\sqrt{\left[\frac{\partial f_1}{\partial x}\right]^2 + \left[\frac{\partial f_1}{\partial y}\right]^2 + 1}}$$

13.9. TEOREMA DE GAUSS.

y

$$\begin{aligned} & \int_{S_1} R(\vec{k} \cdot \vec{n}_1) dS = \\ &= \int_D R(x, y, f_1(x, y)) \frac{-1}{\sqrt{\left[\frac{\partial f_1}{\partial x}\right]^2 + \left[\frac{\partial f_1}{\partial y}\right]^2 + 1}} \sqrt{\left[\frac{\partial f_1}{\partial x}\right]^2 + \left[\frac{\partial f_1}{\partial y}\right]^2 + 1} dA = \\ &= - \int_D R(x, y, f_1(x, y)) dx dy. \end{aligned}$$

De manera análoga, en la cara superior S_2 tenemos

$$\vec{n}_2 \cdot \vec{k} = \frac{1}{\sqrt{\left[\frac{\partial f_2}{\partial x}\right]^2 + \left[\frac{\partial f_2}{\partial y}\right]^2 + 1}}$$

y

$$\int_{S_2} R(\vec{k} \cdot \vec{n}_2) dS = \int_D R(x, y, f_2(x, y)) dx dy.$$

Entonces, tenemos la cadena de igualdades

$$\begin{aligned} \int_{\partial\Omega} R\vec{k} \cdot \vec{n} dS &= \int_{S_1} R\vec{k} \cdot \vec{n}_1 dS + \int_{S_2} R\vec{k} \cdot \vec{n}_2 dS = \\ &= \int_D R(x, y, f_2(x, y)) dx dy - \int_D R(x, y, f_1(x, y)) dx dy = \int_{\Omega} \frac{\partial R}{\partial z} dV. \end{aligned}$$

c.q.d.

Nótese que se puede extender el teorema de Gauss a cualquier región que pueda partirse en regiones del tipo IV. Esto incluye la región entre dos superficies cerradas, una dentro de la otra. La superficie de esta región consta de dos partes orientadas. Se aplica el teorema de la divergencia a dicha región cuando se prueba la ley de Gauss en el teorema 13.17, (enunciado más adelante pero no demostrado).

Ejemplo 1. Considere $\vec{F} = 2x\vec{i} + y^2\vec{j} + z^2\vec{k}$. Sea S la esfera unitaria, $x^2 + y^2 + z^2 = 1$. Evalúa $\int_S \vec{F} \cdot \vec{n} dS$.

Solución:

Por el teorema de Gauss,

$$\int_{\Omega} (\operatorname{div} \vec{F}) dV = \int_S \vec{F} \cdot \vec{n} dS,$$

donde Ω es el volumen acotado por la esfera. La integral de la izquierda es

$$2 \int_{\Omega} (1 + y + z) dV = 2 \int_{\Omega} dV + 2 \int_{\Omega} y dV + 2 \int_{\Omega} z dV.$$

Por simetría, podemos argumentar que las últimas dos integrales son nulas. Así,

$$2 \int_{\Omega} (1 + y + z) dV = 2 \int_{\Omega} dV = 2 \text{Volumen}(bola) = \frac{8}{3} \pi.$$

Los lectores se convencerán de lo difícil de intentar el cálculo directo de $\int_S \vec{F} \cdot \vec{n} dS$

Ejemplo 2. Usa el teorema de la divergencia para evaluar

$$\int_{\partial W} (x^2 + y + z) dS,$$

donde W es la esfera sólida $x^2 + y^2 + z^2 \leq 1$.

Solución:

Para poder aplicar el teorema de la divergencia de Gauss, debemos hallar algún campo vectorial $\vec{F} = F_1 \vec{i} + F_2 \vec{j} + F_3 \vec{k}$ en W con $\vec{F} \cdot \vec{n} = x^2 + y + z$. En cualquier punto $(x, y, z) \in \partial W$, la normal unitaria exterior \vec{n} a ∂W es $\vec{n} = x \vec{i} + y \vec{j} + z \vec{k}$. Por lo tanto, si \vec{F} es el campo vectorial deseado, entonces $\vec{F} \cdot \vec{n} = F_1 x + F_2 y + F_3 z$. Hacemos $F_1 x = x^2$, $F_2 y = y$, $F_3 z = z$ y resolvemos para F_1 , F_2 y F_3 para hallar que

$$\vec{F} = x \vec{i} + \vec{j} + \vec{k},$$

Calculando $\text{div } \vec{F}$ obtenemos $\text{div } \vec{F} = 1$. Así, por el teorema de la divergencia de Gauss,

$$\int_{\partial W} (x^2 + y + z) dS = \int_W dV = \frac{4}{3} \pi.$$

El significado físico de la divergencia es que en un punto P , $\text{div } \vec{F}(P)$ es la tasa del flujo neto hacia el exterior en P por unidad de volumen. Así, si $\text{div } \vec{F}(P) > 0$, consideramos P como una fuente, pues hay un flujo neto hacia el exterior cerca de P . Si $\text{div } \vec{F}(P) < 0$, P se llama sumidero de \vec{F} .

Un campo vectorial \vec{F} de clase C^1 definido en W se llama **solenoidal** (esto es, sin divergencia) si $\text{div } \vec{F}(P) = 0$. Si \vec{F} es solenoidal, tenemos $\int_S \vec{F} \cdot d\vec{S} = 0$ para todas las superficies cerradas S . El recíproco también se puede demostrar rápidamente usando el teorema de Gauss: si $\int_S \vec{F} \cdot d\vec{S} = 0$ para todas las superficies cerradas S , entonces \vec{F} es sin divergencia. Si \vec{F} es sin divergencia, vemos entonces que el flujo de \vec{F} a través de cualquier superficie cerrada S es 0, de modo que si \vec{F} es el campo de velocidad de un fluido, la cantidad neta de fluido que fluye hacia afuera de cualquier región será 0. Así, la misma cantidad de fluido debe fluir hacia adentro de la región que la cantidad que sale (en unidad de tiempo). Por lo tanto, un fluido con esta propiedad se llama **incompresible**.

Teorema 13.17 Ley de Gauss.

Sea M una región en \mathbb{R}^3 del tipo IV. Entonces si $(0, 0, 0) \notin \partial M$, tenemos

$$\int_{\partial M} \frac{\vec{r} \cdot \vec{n}}{r^3} dS = \begin{cases} 4\pi & \text{si } (0, 0, 0) \notin \partial M. \\ 0 & \text{si } (0, 0, 0) \in \partial M. \end{cases}$$

13.9. TEOREMA DE GAUSS.

donde

$$\vec{r}(x, y, z) = x\vec{i} + y\vec{j} + z\vec{k} \quad y \quad r(x, y, z) = \|\vec{r}(x, y, z)\| = \sqrt{x^2 + y^2 + z^2}.$$

Ejercicios.

- 1) Sea $\vec{F} = (x - y)\vec{i} + (y - z)\vec{j} + (z - x)\vec{k}$. Evalúa $\int_{\partial\Omega} \vec{F} \cdot d\vec{S}$, donde Ω es el conjunto de puntos (x, y, z) con $x^2 + y^2 \leq z \leq 1$.
- 2) Sea $\vec{F} = x^2 y\vec{i} + z^8 \vec{j} - 2xy z\vec{k}$. Halla la integral de \vec{F} sobre la superficie del cubo unitario.
- 3) Evalúa la integral $\int \int_S \vec{F} \cdot d\vec{S}$, donde $\vec{F} = x\vec{i} + y\vec{j} - 3\vec{k}$ y S es la superficie de la esfera unitaria.

