

Fascículo 28

Cursos y seminarios de
matemática

Serie A

N. A. Fava, G. Keilhauer, A. R. Larotonda

Cambio de variables,
integrales de superficie y
coordenadas polares en \mathbb{R}^n

Departamento de Matemática

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

2011

Cursos y Seminarios de Matemática – Serie A

Fascículo 28

Comité Editorial:

Carlos Cabrelli (Director)
Departamento de Matemática, FCEyN, Universidad de Buenos Aires.
E-mail: cabrelli@dm.uba.ar

Gabriela Jerónimo
Departamento de Matemática, FCEyN, Universidad de Buenos Aires.
E-mail: jeronimo@dm.uba.ar

Claudia Lederman
Departamento de Matemática, FCEyN, Universidad de Buenos Aires.
E-mail: clerderma@dm.uba.ar

Auxiliar editorial:

Leandro Vendramin
Departamento de Matemática, FCEyN, Universidad de Buenos Aires.
E-mail: lvendramin@dm.uba.ar

ISSN 1853-709X (Versión Electrónica)
ISSN 0524-9643 (Versión Impresa)

Derechos reservados
© 2011 Departamento de Matemática, Facultad de Ciencias Exactas y Naturales,
Universidad de Buenos Aires.

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria – Pabellón I
(1428) Ciudad de Buenos Aires
Argentina.
<http://www.dm.uba.ar>
e-mail. secre@dm.uba.ar
tel/fax: (+54-11)-4576-3335

MEW 4625

CURSOS Y SEMINARIOS DE MATEMATICA

Fascículo 28

N.A.Fava, G.Keilhauer y A.R.Larotonda

Cambio de variables, integrales de
superficie y coordenadas polares
en R^n

1980

19 MAR 2009

BIBLIOTECA
"JULIO REY PASTOR"
Dpto. DE MATEMÁTICA

CSU 517.43

M05625

... ..
... ..
... ..

... ..
... ..

2005 JAN 11

... ..
... ..
... ..

Dedicado al profesor

LUIS ANTONIO SANTALO

en homenaje a sus cuarenta años de
fecunda labor matemática en nuestro
país.

19 MAR 2009

BIBLIOTECA
"JULIO REY PASTOR"
DE DE MATEMATICA

1912

1912

1912

1912

CONTENIDO

PAGINA

CAPITULO I

FORMULA DEL CAMBIO DE VARIABLES

1. Nociones preliminares.....	1
2. Imagen de un conjunto medible por una aplicación lineal.	4
3. Aplicaciones diferenciables.....	10
4. Fórmula del cambio de variables.....	11
5. Coordenadas polares en R^n	22
Ejercicios.....	33

CAPITULO II

LA INTEGRAL SOBRE UNA HIFERSUPERFICIE

1. Hipersuperficies.....	37
2. Espacio tangente a una hipersuperficie.....	43
3. Elemento de área de una hipersuperficie.....	48
4. Integral sobre una hipersuperficie.....	49
5. Expresiones particulares del elemento de área.....	52
6. Resolución de integrales múltiples.....	53
Ejercicios.....	57

CAPITULO III

APLICACIONES

1. El elemento de área de la hipersuperficie S^{n-1}	58
2. Densidades de probabilidad.....	63
3. Algunas densidades útiles en Estadística.....	74

BIBLIOTECA
JULIO REV. PASTOR
DE MATEMÁTICA

Partición de la unidad.....95

[Faint, mostly illegible text, likely bleed-through from the reverse side of the page]

SECRET
MAY 1950

PROLOGO

El presente fascículo está destinado a los alumnos que acaban de terminar un curso de Funciones Reales y necesitan completar su formación con otros temas que, como el que aquí se trata, no suele incluirse en un programa estándar de la materia por razones de tiempo, a pesar de su notoria relevancia.

En concordancia con lo expresado, hemos evitado en el capítulo II el uso de formas diferenciales para definir integración sobre hipersuperficies.

El tema fundamental aquí tratado es la fórmula del cambio de variables para integrales de Lebesgue en \mathbb{R}^n , expuesta en el primer capítulo y que sirve de base para los estudios que se emprenden en los siguientes capítulos; el último de los cuales se ha insertado solamente para ilustrar sobre la utilización del método.

Los autores desean expresar su agradecimiento a la Sra. Silvia C. López por haber llevado a cabo la ardua labor de mecanografiar el manuscrito.

... ..
... ..
... ..
... ..
... ..

... ..
... ..
... ..
... ..
... ..
... ..
... ..

... ..
... ..
... ..

- ERRATA -

- Pág.10, línea 2: debe leerse " $x \in \mathbb{R}^n$ " en lugar de " $x \mathbb{R}^n$ ".
línea 3: en lugar de "aplicación $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ " debe leerse "aplicación lineal $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ".
línea 4: $f(x+h)$ en lugar de $f(x-h)$.
línea 5: "aplicación lineal A" en lugar de "aplicación A".
línea 4↑: Falta el signo \subset entre U y \mathbb{R}^n .
línea 2↑: D_2 debe llevar el superíndice α_2 .
- Pág.15, línea 4: " $h(Q)$ " en lugar de " $h(\varphi)$ ".
línea 5: suprimir "z" en el miembro derecho de (1.15).
línea 10: " $D\varphi(x)$ " en lugar de " $D_\varphi(x)$ ".
línea 13: borrar el paréntesis derecho en el miembro izquierdo.
- Pág.19, línea 6↑: "existen" en lugar de "existe"
- Pág.28, líneas 6 y 7: en lugar de "que es unión finita de subespacios de dimensión $\leq n-2$ " debe decir "contenido en un subespacio de dimensión $n-2$ ".
- Pág.29 En las igualdades (1.21), la integral del segundo miembro se extiende sobre el conjunto $P_+ \times L$.
- Pág.34, línea 9: en lugar de " $f_p * f$ " debe leerse " $f_p * f_q$ ".
- Pág.35, línea 1: " $\Gamma(n+1)$ " en lugar de " $(n+1)$ ".
línea 2: En lugar de " $(2k+1)/2$ " debe leerse " $\Gamma((2k+1)/2)$ ".
- Pág.36, línea 1↑: En lugar de " $\langle Ax, Ax \rangle$ " debe leerse " $\langle Ax, Ax \rangle$ ".
línea 5↑: falta el signo \in entre c y R.
- Pág.37, línea 10: el elemento de la última fila, última columna de la matriz $\varphi'(u)$ es $\partial \varphi_n / \partial u_m$.
línea 13: En lugar de " $J_k(U)$ " debe leerse " $J_k(u)$ ".
- Pág.38 línea 4 : falta " ϵ " entre p y H.

Pág. 39, línea 7: $g(u, 0) = \varphi(u)$ en lugar de " $g(u, 0) = (u)$ ".

Pág. 43, línea 10: Debe leerse "definimos" en lugar de "ponemos".

Pág. 48, línea 11↑: La ecuación poco visible es $\varphi(U) = \psi(V)$.
línea 1↑: "a" en lugar de "da".

Pág. 51, línea 8↑: $(k_j, j \in J)$ es poco visible.

Pág. 55, línea 6: Debe leerse "una representación paramétrica o carta local de la hipersuperficie...".

línea 6↑: "válida" en lugar de "vál da".

línea 3↑: falta f en el numerador del integrando, del miembro derecho.

Pág. 62 En el miembro derecho de (3.9) al numerador le falta el exponente $n/2$.

Pág. 65, línea 6: $\left| \frac{\partial x}{\partial y} \right|$ en lugar de " $\left| \frac{\partial x}{\partial y} \right|'$ ".

SECRET
NO FORN DISSEM

FORMULA DEL CAMBIO DE VARIABLES PARA INTEGRALES DE LEBESGUE
EN R^n

1. Nociones preliminares

Entre éstas necesitamos la noción de norma de una aplicación lineal, que vamos a definir a continuación.

Sean X e Y dos espacios normados reales o complejos y sea $T: X \rightarrow Y$ una aplicación lineal de X en Y . Comenzaremos demostrando la equivalencia de las siguientes afirmaciones:

(a) T es continua

(b) T es continua en el punto $0 \in X$

(c) existe un número real C tal que $\|Tx\| \leq C \|x\|$ para cada vector $x \in X$

(d) T es uniformemente continua.

Claramente (a) implica (b). Supongamos que T es continua en el punto $0 \in X$. Puesto que $T0 = 0$, existe un número positivo δ tal que $\|Tx\| \leq 1$ a condición de que se verifique $\|x\| \leq \delta$. Si x es un vector no nulo del espacio X , el vector $\delta x / \|x\|$ tiene norma igual a δ y por consiguiente, $\|T(\delta x / \|x\|)\| \leq 1$, de donde $\|Tx\| \leq \delta^{-1} \|x\|$. Como esta igualdad se cumple de manera obvia si $x = 0$, hemos demostrado que (b) implica (c).

Puesto que T es lineal, si (c) es verdadera, entonces tendremos

$$\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \leq C \|x_1 - x_2\|, \text{ de donde resulta (d).}$$

Finalmente, es obvio que (d) implica (a).

Si T es continua, entre los números C que verifican la relación

$$(1.1) \quad \|Tx\| \leq C \|x\| \quad (\text{para cada } x \in X)$$

hay uno mínimo. En efecto, si C_0 es el ínfimo del conjunto formado por todos los números C que verifican (1.1), entonces existe una sucesión $(C_n, n = 1, 2, \dots)$ de tales números, tal que $C_n \rightarrow C_0$ cuando $n \rightarrow \infty$. Puesto que para cada n y para cada x , $\|Tx\| \leq C_n \|x\|$, pasando al límite resulta $\|Tx\| \leq C_0 \|x\|$.

El número C_0 , es decir, la mínima constante C que verifica (1.1), se llama la norma de la aplicación lineal T y se denota por $\|T\|$; de modo que

$$(1.2) \quad \|Tx\| \leq \|T\| \|x\| \quad (x \in X)$$

Para calcular la norma de una aplicación lineal continua T se tiene la fórmula

$$(1.3) \quad \|T\| = \sup_{\|x\|=1} \|Tx\|,$$

cuya demostración es una consecuencia inmediata de la definición y del hecho de que si x es un vector no nulo, entonces el vector $x/\|x\|$ tiene norma igual a uno.

Llamando K al cuerpo de escalares de los espacios X e Y , el conjunto $L(X, Y)$ formado por todas las aplicaciones lineales continuas de X en Y es un K -espacio vectorial si se adoptan las definiciones habituales.

$$(1.4) \quad (T_1 + T_2)x = T_1x + T_2x, \quad (aT)x = aTx \quad (a \in K).$$

Más aún, $L(X, Y)$ es un espacio normado si se define la norma de cada uno de sus elementos por medio de la fórmula (1.3), cosa que haremos siempre desde ahora en adelante. Por razones obvias,

nos limitamos a enunciar y a sugerir las demostraciones de los hechos que son esenciales para el ulterior desarrollo del tema que nos interesa.

Ejercicios

1. Probar las relaciones $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$, $\|aT\| = |a| \|T\|$ y verificar que $\|T\| = 0$ si y sólo si $T = 0$.
2. Concluir del ejercicio precedente que $|\|T_1\| - \|T_2\|| \leq \|T_1 - T_2\|$.

En el caso particular en que se verifica $X=Y$, además de las operaciones vectoriales (1.4) podemos introducir en el espacio normado $E(X) = L(X, X)$ la noción de producto o composición, por medio de la fórmula

$$(1.5) \quad (T_1 T_2)x = T_1(T_2 x);$$

y es fácil probar que se verifica la desigualdad muy útil

$$(1.6) \quad \|T_1 T_2\| \leq \|T_1\| \|T_2\|.$$

En efecto, $\|(T_1 T_2)x\| = \|T_1(T_2 x)\| \leq \|T_1\| \|T_2 x\| \leq \|T_1\| \|T_2\| \|x\|$, de donde resulta (1.6) en virtud de la definición de la norma de una aplicación lineal.

Si $x = (x_1, \dots, x_n)$ es un vector del espacio R^n , la norma euclidiana o módulo del vector x es el número $|x| = (x_1^2 + \dots + x_n^2)^{1/2}$.

Si T es una aplicación lineal de R^n en sí mismo cuya matriz en la base canónica de R^n es (a_{ik}) , entonces la relación $y = Tx$ es equivalente a las ecuaciones

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (1 \leq i \leq n),$$

y es fácil probar que en tal caso, $|Tx| = |y| < C|x|$, donde C representa la raíz cuadrada de la suma de los cuadrados de todos los elementos de la matriz de T .

(1.7) COROLARIO. Toda aplicación lineal T de \mathbb{R}^n en sí mismo es continua. Si T es invertible, entonces es un homeomorfismo.

En particular, si T es invertible, entonces T aplica conjuntos abiertos en conjuntos abiertos y conjuntos de clase (G_δ) en conjuntos de la misma clase.

2. Imagen de un conjunto medible por una aplicación lineal

En esta sección estudiaremos la forma en que actúa una aplicación lineal T de \mathbb{R}^n en sí mismo sobre un conjunto medible $E \subset \mathbb{R}^n$. Al hablar de "la matriz" de T nos referimos exclusivamente a la matriz de T en la base canónica de \mathbb{R}^n .

Si (a_{ik}) es la matriz de T e $y = Tx$, la relación entre las coordenadas de los vectores x e y se puede escribir en la forma de un producto matricial

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o bien en la forma de un sistema de ecuaciones

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (1 \leq i \leq n).$$

Necesitaremos considerar tres tipos especiales de aplicaciones

lineales de R^n a las cuales llamaremos aplicaciones elementales, a saber,

$$T_a(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = (x_1, \dots, x_j, \dots, x_i, \dots, x_n)$$

$$T_b(x_1, \dots, x_i, \dots, x_n) = (x_1, \dots, \lambda x_i, \dots, x_n) \quad (\lambda \neq 0)$$

$$T_c(x_1, \dots, x_i, \dots, x_n) = (x_1, \dots, x_i + x_j, \dots, x_n) \quad (i \neq j)$$

Las matrices que corresponden a estas aplicaciones se llaman matrices elementales.

Es evidente que la inversa de cada aplicación elemental es una aplicación del mismo tipo. Además, calculando las matrices de estas aplicaciones se verifica fácilmente que $\det T_a = -1$, $\det T_b = \lambda$, $\det T_c = 1$.

(1.8) TEOREMA. Toda aplicación lineal invertible T de R^n en sí mismo es un producto (finito) de aplicaciones elementales.

Daremos solamente un esbozo de la demostración que suele estudiarse en los cursos de Algebra Lineal.

Recordemos que cualquier matriz no singular (a_{ik}) se puede transformar en la matriz unitaria I_n al cabo de un número finito de operaciones de fila, a saber,

- 1°) permutar dos filas de la matriz
- 2°) multiplicar una fila por un coeficiente no nulo
- 3°) sumar una fila a otra fila.

Por otra parte, cualquier operación de fila sobre una matriz se puede realizar multiplicando a izquierda la matriz dada por la

correspondiente matriz elemental.

Denotando por $A = (a_{ik})$ la matriz de T , tendremos $\det A \neq 0$ y por consiguiente existen matrices elementales E_1, E_2, \dots, E_k tales que $E_k \dots E_2 E_1 A = I_n$, de donde resulta $A = (E_k \dots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \dots E_k^{-1}$ y el teorema queda demostrado.

En adelante, el símbolo $|E|$ se usará para denotar la medida de Lebesgue del conjunto $E \subset \mathbb{R}^n$.

(1.9) TEOREMA. Si T es una aplicación lineal de \mathbb{R}^n en sí mismo, entonces para cualquier conjunto medible $E \subset \mathbb{R}^n$, la imagen $T(E)$ es medible y además, $|T(E)| = |\det T| |E|$.

Si T es singular, entonces $\det T = 0$ y $T(\mathbb{R}^n)$ es un subespacio propio de \mathbb{R}^n , lo cual implica que el conjunto $T(E) \subset T(\mathbb{R}^n)$ tiene medida nula, y el teorema es trivial en este caso. Suponiendo, pues, que T es invertible, comenzaremos verificando las afirmaciones del enunciado en el caso particular en que T es una aplicación elemental, en tanto que $E = I$ es un intervalo de \mathbb{R}^n :

$$I = \{x \in \mathbb{R}^n : a_i < x_i \leq b_i \quad (i=1, 2, \dots, n)\}.$$

La aplicación elemental T_a transforma el intervalo I en un intervalo de la misma medida, mientras que T_b transforma I en un intervalo cuyos lados coinciden con los de I , con la excepción de uno solo de ellos que se transforma según una homotecia de razón $|\lambda|$, lo que hace que la verificación sea muy fácil en el caso de estas aplicaciones. En cuanto a una aplicación elemental del tipo T_c estudiaremos, para fijar ideas, la aplicación T definida por

$$T(x_1, \dots, x_n) = (x_1, \dots, x_{n-1}, x_1 + x_n),$$

de modo que si $y = Tx$, entonces tendremos $y_1 = x_1, \dots, y_{n-1} = x_{n-1}, y_n = x_1 + x_n$.

La inversa de la aplicación T se expresa entonces por medio de las ecuaciones $x_1 = y_1, \dots, x_{n-1} = y_{n-1}, x_n = y_n - y_1$, de las que se infiere inmediatamente que

$$T(I) = \{y \in \mathbb{R}^n : a_1 < y_1 \leq b_1, \dots, a_{n-1} < y_{n-1} \leq b_{n-1}, a_n + y_1 < y_n \leq b_n + y_1\}.$$

Por consiguiente, si $a_1 < y_1 \leq b_1$, la sección

$$T(I)^{y_1} = \{(y_2, \dots, y_n) \in \mathbb{R}^{n-1} : (y_1, y_2, \dots, y_n) \in T(I)\}$$

es igual al intervalo definido por las desigualdades $a_2 < y_2 \leq b_2, \dots, a_{n-1} < y_{n-1} \leq b_{n-1}, a_n + y_1 < y_n \leq b_n + y_1$. Es claro también que la sección es vacía si no se verifican las desigualdades $a_1 < y_1 \leq b_1$.

Puesto que I es un conjunto de clase (G_δ) no hay duda de que T(I) es medible, y el Teorema de Fubini nos permite calcular su medida:

$$|T(I)| = \int_{a_1}^{b_1} |T(I)^{y_1}| dy_1 = (b_2 - a_2) \dots (b_{n-1} - a_{n-1}) (b_n - a_n) = |I|.$$

Puesto que $\det T = 1$ en el caso bajo consideración, queda demostrado que si T es una aplicación elemental, entonces para cualquier intervalo I de \mathbb{R}^n se verifica $|T(I)| = |\det T| |I|$.

Manteniendo la hipótesis de que T sea una aplicación elemental, supongamos sucesivamente:

BIBLIOTECA
"JULIO REY PASTOR"
Dpto. DE MATEMATICA

1°) E es un conjunto abierto. Entonces existe una sucesión $(I_k, k=1,2,\dots)$ de intervalos disjuntos cuya unión es E y por consiguiente $T(E)$, que es la unión de los conjuntos disjuntos $T(I_k)$, tiene medida

$$|T(E)| = \sum_{k=1}^{\infty} |T(I_k)| = |\det T| \sum_{k=1}^{\infty} |I_k| = |\det T| \cdot |E|.$$

2°) E es un conjunto de clase (G_δ) acotado. Entonces existe una sucesión decreciente de conjuntos abiertos acotados $G_1 \supset G_2 \supset \dots \supset G_k \supset \dots$ cuya intersección es E. Puesto que $T(E)$ es la intersección de la sucesión decreciente formada por los conjuntos abiertos acotados $T(G_k)$, se tiene

$$|T(E)| = \lim_{k \rightarrow \infty} |T(G_k)| = |\det T| \lim_{k \rightarrow \infty} |G_k| = |\det T| |E|.$$

3°) E es un conjunto de clase (G_δ) . Llamando E_k a la intersección de E con la bola abierta $B(0,k)$ con centro en el origen y radio k, es claro que los conjuntos E_k forman una sucesión creciente cuya unión es E y que cada uno de ellos verifica las condiciones de la etapa anterior. Por consiguiente,

$$|T(E)| = \lim_{k \rightarrow \infty} |T(E_k)| = |\det T| \lim_{k \rightarrow \infty} |E_k| = |\det T| |E|.$$

4°) E es un conjunto de medida nula. Entonces existe un conjunto D de clase (G_δ) tal que $E \subset D$ y $|D| = 0$. Puesto que $|T(D)| = |\det T| |D| = 0$, se deduce que el conjunto $T(E) \subset T(D)$ es de medida nula, es decir, T aplica conjuntos de medida nula en conjuntos de medida nula.

5°) E es un conjunto medible arbitrario. Entonces existen un conjunto D de clase (G_δ) y un conjunto N de medida nula, tales que $E = D - N$. Por lo tanto $T(E) = T(D) - T(N)$ es medible, y además, $|T(E)| = |T(D)| = |\det T| |D| = |\det T| |E|$. Hemos verificado las afirmaciones del teorema en el caso particular de que T sea una aplicación elemental. Ahora bien; si T es una aplicación lineal invertible de R^n en sí mismo, entonces existen aplicaciones elementales T_1, \dots, T_k tales que $T = T_1 T_2 \dots T_k$ y la propiedad multiplicativa del determinante, juntamente con inducción sobre k, nos permiten completar fácilmente la demostración. Q.E.D.

(1.10) COROLARIO. Si $T: R^n \rightarrow R^n$ es un aplicación ortogonal, entonces para cualquier conjunto medible $E \subset R^n$, la imagen $T(E)$ es medible y tiene medida igual a la medida de E.

Ejercicio

Demostrar que si v_1, \dots, v_n son vectores de R^n , entonces el paralelepípedo

$$P = \{x: x = \sum_{k=1}^n t_k v_k \quad (0 \leq t_k \leq 1)\}$$

tiene medida

$$|P| = \left| \det \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right|$$

Pista: la aplicación lineal T definida por $T(t_1, \dots, t_n) = t_1 v_1 + \dots + t_n v_n$ transforma el cubo unitario de R^n en el conjunto P.

$$T(t_1, \dots, t_n) = \sum_{k=1}^n t_k v_k$$

3. Aplicaciones Diferenciales

Una aplicación $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en el punto $x \in \mathbb{R}^n$ si existe una aplicación $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que

$$\lim_{h \rightarrow 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0.$$

La única aplicación A que verifica esta relación se llama la derivada de f en x y se denota por $Df(x)$.

Si $y = f(x)$ podemos escribir

$$y_i = f_i(x_1, \dots, x_n) \quad (1 \leq i \leq m),$$

y la matriz de A con respecto a las bases canónicas de \mathbb{R}^n y \mathbb{R}^m resulta ser la matriz de m filas por n columnas

$$f'(x) = \left(\frac{\partial y_i}{\partial x_k} \right) = (D_k f_i(x)) \quad (1 \leq i \leq m, 1 \leq k \leq n),$$

donde $D_k f_i(x) = \partial f_i / \partial x_k$. La matriz $f'(x)$ se llama matriz jacobiana de f en x .

En el caso en que $m = n$, el número $\det f'(x)$ se llama el determinante jacobiano de f en x y se representa por las notaciones usuales

$$\frac{\partial y}{\partial x} = \frac{\partial (y_1, \dots, y_n)}{\partial (x_1, \dots, x_n)} = \frac{\partial (f_1, \dots, f_n)}{\partial (x_1, \dots, x_n)}.$$

Las definiciones precedentes mantienen su sentido si el dominio de f es un conjunto abierto $U \subset \mathbb{R}^n$. En tal caso, decimos que f es diferenciable en U si es diferenciable en cada punto de U y decimos que f es de clase C^m en U si las derivadas parciales $D_1^{\alpha_1} D_2^{\alpha_2} \dots D_n^{\alpha_n} f_i$ existen y son continuas en U siempre que $\alpha_1 + \alpha_2 + \dots + \alpha_n \leq m$. Por lo

general trabajaremos con funciones de clase C^1 .

Toda función de clase C^1 en un conjunto abierto U es diferenciable en dicho conjunto.

Un hecho básico es la llamada regla de la cadena, la cual establece que si $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es diferenciable en el punto x , y $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$ es diferenciable en $f(x)$, entonces la composición $g \circ f$ es diferenciable en x , y además, $D(g \circ f)(x) = Dg(f(x)) \circ Df(x)$.

Sean U y V dos conjuntos abiertos de \mathbb{R}^n . Una aplicación biyectiva $h: U \rightarrow V$ se llama un difeomorfismo si es diferenciable y si la aplicación inversa $h^{-1}: V \rightarrow U$ es también diferenciable.

Mencionemos, por último, que si $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una aplicación lineal, entonces en cualquier punto x de \mathbb{R}^n se cumple $Df(x) = f$, como se comprueba inmediatamente a partir de la definición.

Para mayores detalles, incluyendo las demostraciones de los hechos mencionados, recomendamos consultar la obra de M. Spivak [7] cuyas notaciones hemos adoptado.

4. Fórmula del cambio de variables

En esta sección resultará muy cómodo introducir en el espacio \mathbb{R}^n formado por las n -uplas $x = (x_1, \dots, x_n)$ de números reales una norma distinta de la norma euclidiana, a saber,

$$(1.11) \quad \|x\| = \max_{1 \leq i \leq n} |x_i|$$

Comenzaremos demostrando que si (a_{ik}) representa la matriz de una aplicación lineal A de \mathbb{R}^n en sí mismo, entonces la norma de A

con respecto a la norma (1.11) se calcula por medio de la fórmula

$$(1.12) \quad \|A\| = \sup_{\|x\|=1} \|Ax\| = \max_{1 \leq i \leq n} \sum_{k=1}^n |a_{ik}|.$$

En efecto, si $y = Ax$, entonces llamando C al último miembro de (1.12) tendremos:

$$\|y\| = \left| \sum_{k=1}^n a_{ik} x_k \right| \leq \sum_{k=1}^n |a_{ik}| \|x\| \leq C \|x\|.$$

Luego $\|Ax\| = \|y\| \leq C \|x\|$ y por consiguiente, $\|A\| \leq C$. Por otro lado, si m es un entero comprendido entre 1 y n que verifica

$$\sum_{k=1}^n |a_{mk}| = C, \text{ definimos un vector } x = (x_1, \dots, x_n) \text{ poniendo}$$

$$x_k = 1 \text{ si } a_{mk} \geq 0, \quad x_k = -1 \text{ en caso contrario.}$$

Entonces $\|x\| = 1$ y por lo tanto,

$$\|A\| \geq \|Ax\| \geq \left| \sum_{k=1}^n a_{mk} x_k \right| = \sum_{k=1}^n |a_{mk}| = C.$$

Las desigualdades obtenidas permiten concluir que $\|A\| = C$, como queríamos probar.

En todo lo que sigue de esta sección, supondremos permanentemente que O_1 y O_2 son dos conjuntos abiertos de \mathbb{R}^n y que $h: O_1 \rightarrow O_2$ es una aplicación biyectiva de clase C^1 con la propiedad de que la aplicación lineal $j(x) = Dh(x)$ es invertible en cada punto $x \in O_1$, lo cual equivale a afirmar que el determinante jacobiano $\det j(x) = \det h'(x)$ es distinto de cero en cada punto de O_1 .

El teorema de la función inversa afirma que, en estas condiciones, la aplicación inversa $h^{-1}: O_2 \rightarrow O_1$ es también de clase

C^1 .

Para cada punto $y \in O_2$ pondremos $j^{-1}(y) = Dh^{-1}(y)$ e introduci-
mos las notaciones $J(x) = |\det j(x)|$, $J^{-1}(y) = |\det j^{-1}(y)|$.

Aplicando la regla de la cadena en la relación $h^{-1} \circ h = id_{O_1}$,
(en general, id_A denotará la aplicación idéntica de A en sí
mismo) obtenemos $Dh^{-1}(h(x)) \circ Dh(x) = id_{R^n}$, de modo que
si $y = h(x)$, entonces tendremos

$$j^{-1}(y) \circ j(x) = id_{R^n}.$$

Tomando determinantes en esta relación, resulta $J^{-1}(y) J(x) = 1$,
siempre que $y = h(x)$.

Ahara estamos en condiciones de enunciar y probar el teorema
más importante de estas notas.

(1.13) TEOREMA (fórmula del cambio de variables).

Si $f: O_2 \rightarrow R^n$ es una función medible no negativa sobre O_2 ,
entonces la composición $f \circ h$ es medible sobre O_1 , y, a-
demás,

$$\int_{O_2} f(y) dy = \int_{O_1} f \circ h(x) J(x) dx.$$

La demostración se basa en el siguiente lema.

(1.14) LEMA. Si Q es un cubo de R^n que verifica $\bar{Q} \subset O_1$, entonces

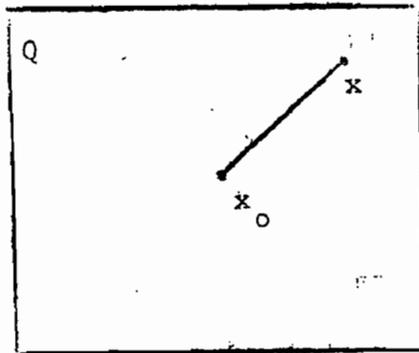
$$|h(Q)| \leq \int_Q J(x) dx.$$

Puesto que $h(Q) \subset h(\bar{Q})$, podemos suponer que Q es compacto,
digamos

$$Q = \{x : \|x - x_0\| \leq \lambda\}.$$

El punto $x_0 = (x_{01}, x_{02}, \dots, x_{0n})$ es el centro de Q , cuyos lados tienen longitud igual a 2λ . (figura 1).

FIGURA 1



En virtud del teorema del valor medio, para cualquier $x \in Q$

tenemos

$$h_i(x) - h_i(x_0) = \sum_{k=1}^n D_k h_i(\bar{x}_i) (x_k - x_{0k}),$$

donde $\bar{x}_i = x_0 + \theta_i(x - x_0)$, $0 < \theta_i < 1$, de modo que $\bar{x}_i \in Q$ (cada cubo es un conjunto convexo). Luego

$$|h_i(x) - h_i(x_0)| \leq \sum_{k=1}^n |D_k h_i(\bar{x}_i)| |x_k - x_{0k}| \leq$$

$$\leq \lambda \sum_{k=1}^n |D_k h_i(\bar{x}_i)| \leq \lambda \|j(\bar{x}_i)\| \leq$$

$$\leq \lambda \max_{x \in Q} \|j(x)\|, \text{ es decir}$$

$$\|h(x) - h(x_0)\| \leq \lambda \max_{x \in Q} \|j(x)\|.$$

Esto muestra que $h(Q)$ está contenido en el cubo Q^* formado por todos los puntos y que verifican $\|y - h(x_0)\| \leq \lambda \max_{x \in Q} \|j(x)\|$, de donde resulta

$$|h(Q)| \leq |Q^*| \leq (2\lambda)^n \left\{ \max_{x \in Q} \|j(x)\| \right\}^n \text{ es decir}$$

$$(1.15) \quad |h(Q)| \leq 2^n \left\{ \max_{x \in Q} \|j(x)\| \right\}^n |Q|.$$

Para cualquier aplicación lineal invertible A de \mathbb{R}^n en sí mismo, pongamos

$$(1.16) \quad \psi = A^{-1} \circ h.$$

La regla de la cadena nos da

$$j_\psi(x) = D_\psi(x) = A^{-1} \circ j(x) \quad (x \in Q_1),$$

y aplicando (1.15) a la función (1.16), con j_ψ en lugar de j , obtenemos

$$|A^{-1}h(Q)| \leq \left\{ \max_{x \in Q} \|A^{-1}j(x)\| \right\}^n |Q|$$

y en virtud del teorema (1.9), resulta

$$(1.17) \quad |h(Q)| \leq |\det A| \left\{ \max_{x \in Q} \|A^{-1}j(x)\| \right\}^n |Q|$$

para cualquier cubo cerrado $Q \subset \mathbb{R}^n$ y cualquier aplicación lineal invertible $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Pongamos ahora

$$M = \max_{y \in h(Q)} \|j^{-1}(y)\|.$$

Puesto que Q es compacto y h es de clase C^1 , dado $\varepsilon > 0$ existe $\delta > 0$, tal que $u \in Q, v \in Q, \|u-v\| < \delta$ implica

$$\|j(u) - j(v)\| < \varepsilon.$$

Dividamos el cubo Q en cubos cerrados no rampantes Q_1, Q_2, \dots, Q_m de diámetro menor que cierto número positivo $\eta \leq \delta$. Si x_k es el centro del cubo Q_k , pongamos $y_k = h(x_k)$ ($k=1, 2, \dots, m$) y apliquemos (1.17) al cubo Q_k con $A = j(x_k)$, de manera que $A^{-1} = j^{-1}(y_k)$ y por lo tanto,

$$|h(Q_k)| \leq J(x_k) \left\{ \max_{x \in Q_k} \|j^{-1}(y_k) j(x)\| \right\}^n |Q_k|.$$

Denotando por E la aplicación idéntica de R^n , para cada $x \in Q_k$ tenemos

$$\|j^{-1}(y_k) j(x)\| - 1 = \|j^{-1}(y_k) j(x)\| - \|E\|$$

$$\leq \|j^{-1}(y_k) j(x) - E\|$$

$$= \|j^{-1}(y_k) \{j(x) - j(x_k)\}\| \leq$$

$$\|j^{-1}(y_k)\| \cdot \|j(x) - j(x_k)\| < M\varepsilon.$$

Por consiguiente,

$$\|j^{-1}(y_k) j(x)\| < 1 + M\varepsilon \quad (x \in Q_k), \quad (1.18)$$

de donde

$$|h(Q_k)| \leq (1 + M\varepsilon)^n J(x_k) |Q_k|.$$

Luego,

$$|h(Q)| \leq \sum_{k=1}^m |h(Q_k)| \leq$$

$$\leq (1 + M\epsilon)^n \sum_{k=1}^m J(x_k) |Q_k|,$$
 y haciendo $\eta \rightarrow 0$, obtenemos

$$|h(Q)| \leq (1 + M\epsilon)^n \int_Q J(x) dx.$$

Puesto que ϵ es arbitrario, el lema queda demostrado.

Obsérvese que puesto que h es un homeomorfismo, la imagen por h de cualquier conjunto de clase (G_δ) contenido en O_1 es un conjunto de clase (G_δ) contenido en O_2 . Por esta razón la imagen $h(Q)$ es medible aunque Q no sea cerrado, pues cada cubo es un conjunto de clase (G_δ) .

(1.15) LEMA. Si $E \subset O_1$ es un conjunto medible, entonces $h(E)$ es medible, y además,

$$|h(E)| \leq \int_E J(x) dx.$$

La demostración de este lema se realiza por etapas. Supongamos sucesivamente:

1°) E es abierto. Entonces existe una sucesión de cubos disjuntos Q_k ($k=1,2,\dots$) tales que

$$E = \bigcup_{k=1}^{\infty} Q_k, \quad \bar{Q}_k \subset O_1.$$

Por consiguiente,

$$\begin{aligned} |h(E)| &= \sum_{k=1}^{\infty} |h(Q_k)| \leq \sum_{k=1}^{\infty} \int_{Q_k} J(x) dx = \\ &= \int_E J(x) dx. \end{aligned}$$

2°) E es un conjunto de clase (G_δ) acotado y a distancia positiva del complemento de O_1 :

$$\text{dist}(E, O_1^c) = \delta > 0$$

Sean U y K, respectivamente, los conjuntos de puntos x definidos por las relaciones

$$\text{dist}(x, E) < \frac{\delta}{2}, \quad \text{dist}(x, E) \leq \frac{\delta}{2}$$

Luego, $E \subset U \subset K \subset O_1$ y existe una sucesión de conjuntos abiertos $G_i (i=1, 2, \dots)$ que verifican

$$U \supset G_1 \supset G_2 \supset \dots, \quad E = \bigcap_{i=1}^{\infty} G_i$$

Por consiguiente

$h(K) \supset h(G_1) \supset h(G_2) \supset \dots$, $h(E) = \bigcap_{i=1}^{\infty} h(G_i)$, y en virtud de la etapa anterior,

$$|h(G_i)| \leq \int_{G_i} J(x) dx.$$

Puesto que K es compacto, podemos pasar al límite haciendo $i \rightarrow \infty$, con lo cual resulta

$$\begin{aligned} |h(E)| &= \lim_{i \rightarrow \infty} |h(G_i)| \leq \lim_{i \rightarrow \infty} \int_{G_i} J(x) dx = \\ &= \int_E J(x) dx \end{aligned}$$

3°) E es un conjunto de clase (G_δ) . Consideremos la sucesión de conjuntos

$$E_k = \{x \in E : |x| < k, \text{dist}(x, O_1^c) > \frac{1}{k}\}$$

Cada E_k es un conjunto de clase (G_δ) acotado y situado

a distancia positiva del complemento de O_1 . Además,

$$E_1 \subset E_2 \subset \dots, \quad E = \bigcup_{k=1}^{\infty} E_k$$

y en virtud de la etapa precedente,

$$|h(E_k)| \leq \int_{E_k} J(x) dx.$$

Luego,

$$\begin{aligned} |h(E)| &= \lim_{k \rightarrow \infty} |h(E_k)| \leq \lim_{k \rightarrow \infty} \int_{E_k} J(x) dx = \\ &= \int_E J(x) dx. \end{aligned}$$

4°) E es un conjunto de medida nula. Entonces existe un conjunto D de clase (G_δ) , tal que $E \subset D \subset O_1$, $|D| = 0$,

de donde

$$|h(D)| \leq \int_D J(x) dx = 0$$

Puesto que $h(E) \subset h(D)$, vemos que $h(E)$ es un conjunto de medida nula.

Hemos demostrado que h aplica cada subconjunto de medida nula de O_1 en un subconjunto de medida nula de O_2 .

5°) E es un conjunto medible arbitrario contenido en O_1 .

Entonces existe un conjunto D de clase (G_δ) y un conjunto N de medida nula, tales que $E = D - N$. Luego

$h(E) = h(D) - h(N)$ es medible por ser la diferencia

entre un conjunto de clase (G_δ) y un conjunto de medida nula. Además,

$$|h(E)| = |h(D)| \leq \int_{D \cap E} J(x) dx = \int_E J(x) dx$$

y el lema queda demostrado.

La primera consecuencia de (1.15) es que si $f: O_2 \rightarrow \mathbb{R}$ es medible, entonces $f \circ h$ es medible sobre O_1 . En efecto, si B es un conjunto boreliano de la recta real, entonces $(f \circ h)^{-1}(B) = h^{-1}(f^{-1}(B))$ es un subconjunto medible de O_1 , pues h^{-1} verifica con respecto a O_2 las mismas propiedades que verifica h con respecto a O_1 y en consecuencia, h^{-1} aplica cada subconjunto medible de O_2 en un subconjunto medible de O_1 .

Vamos a probar que si $f: O_2 \rightarrow \mathbb{R}$ es una función medible no negativa, entonces

$$(1.16) \int_{O_2} f(y) dy \leq \int_{O_1} f(h(x)) J(x) dx.$$

Si $f = \chi_M$ es la función característica de un conjunto medible $M \subset O_2$, poniendo $E = h^{-1}(M)$ tendremos

$$\int_{O_2} \chi_M(y) dy = |h(E)| \leq \int_{O_1} J(x) dx =$$

$$= \int_{O_1} \chi_E(x) J(x) dx = \int_{O_1} \chi_M(h(x)) J(x) dx.$$

Por la linealidad de la integral, la desigualdad obtenida se extiende al caso en que f es una función simple no negativa; y en el caso general, existe una sucesión creciente de funciones simples no negativas que converge puntualmente a f y la desigualdad (1.16) resulta pasando al límite, en virtud del teorema de

Beppo-Levi.

Si ahora aplicamos (1.16) a la función $g(x) = f(h(x)) J(x)$ sobre el conjunto O_1 , con h^{-1} en lugar de h , obtenemos:

$$\begin{aligned} \int_{O_1} g(x) dx &\leq \int_{O_2} g(h^{-1}(y)) J^{-1}(y) dy = \\ &= \int_{O_2} f(y) J(h^{-1}(y)) J^{-1}(y) dy = \\ &= \int_{O_2} f(y) dy. \end{aligned}$$

El teorema (1.13) queda así completamente demostrado.

Si $f: O_2 \rightarrow R$ es una función medible con valores de distinto signo, recurriendo a la descomposición $f = f^+ - f^-$ de la función f como diferencia de dos funciones medibles no negativas, obtenemos el siguiente corolario:

(1.17) COROLARIO. La función medible $f(y)$ es integrable sobre O_2 si y sólo si $f \circ h(x) J(x)$ es integrable sobre O_1 , y en tal caso,

$$\int_{O_2} f(y) dy = \int_{O_1} f \circ h(x) J(x) dx.$$

Recordando que $J(x) = \left| \frac{\partial y}{\partial x} \right|$, la última fórmula se puede escribir en la forma más sugestiva.

$$\int_{O_2} f(y) dy = \int_{O_1} f(h(x)) \left| \frac{\partial y}{\partial x} \right| dx$$

en notoria analogía con la fórmula del cambio de variables para funciones definidas sobre un intervalo de la recta, de la cual

la fórmula que hemos demostrado representa una muy amplia generalización (obsérvese, en particular, que en ningún momento hemos supuesto que los conjuntos O_1 y O_2 sean acotados). La demostración que acabamos de exponer, basada en el lema (1.14) se debe a Jacob Schwartz [4] :

5. Coordenadas polares en R^n .

Las fórmulas que relacionan las coordenadas cartesianas de un punto $x = (x_1, x_2)$ del plano R^2 con sus correspondientes coordenadas polares r y θ son:

$$(1.18) \quad x_1 = r \cos \theta, \quad x_2 = r \operatorname{sen} \theta$$

donde $r > 0$, $0 \leq \theta < 2\pi$. Los números r y θ están unívocamente determinados dentro de sus respectivos intervalos por las relaciones

$$r = \sqrt{x_1^2 + x_2^2}, \quad \cos \theta = \frac{x_1}{r}, \quad \operatorname{sen} \theta = \frac{x_2}{r}$$

Las ecuaciones (1.18) definen una aplicación biyectiva del producto cartesiano o "franja" $(0, \infty) \times [0, 2\pi]$ (figura 2) sobre el conjunto $R^2 - \{0\}$ que consiste de todo el plano menos el origen de coordenadas.

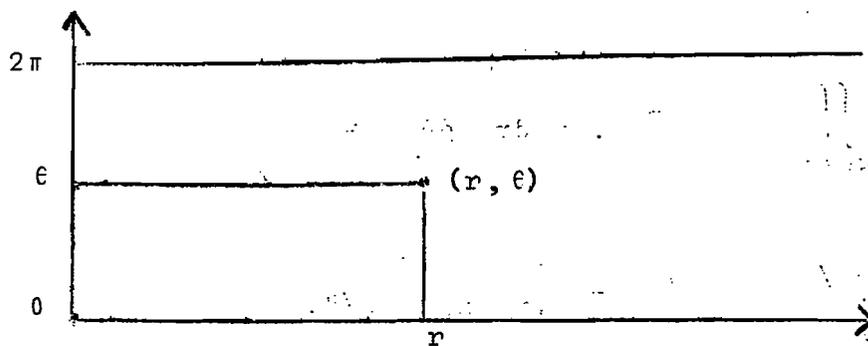


FIGURA 2

Para tener una correspondencia biyectiva entre conjuntos abiertos habría que eliminar de dicha franja la semirrecta $\theta = 0$, lo cual equivale a eliminar del plano \mathbb{R}^2 no sólo el origen sino toda la semirrecta $x_1 \geq 0, x_2 = 0$. Los conjuntos eliminados tienen medida plana nula y por este motivo no tienen ninguna influencia en el valor de las integrales con respecto a la medida de Lebesgue en el plano, de manera que incluirlos o no es irrelevante desde ese punto de vista.

Como es fácil de verificar, el jacobiano de la aplicación (1.18) tiene el valor

$$\frac{\partial(x_1, x_2)}{\partial(r, \theta)} = r.$$

Para ilustrar todo lo dicho, es bueno analizar el conocido ejemplo de la función

$$f(x_1, x_2) = e^{-(x_1^2 + x_2^2)},$$

cuya integral sobre el cuadrante $x_1 > 0, x_2 > 0$, de acuerdo con el teorema de Fubini y el teorema (1.13), es

$$\int_0^{\infty} dx_1 \int_0^{\infty} dx_2 e^{-(x_1^2 + x_2^2)} =$$

$$\begin{aligned}
 &= \int_{0 < \theta < \pi/2} \int_0^{\infty} e^{-r^2} \cdot r \, dr \, d\theta = \\
 &= \int_0^{\pi/2} d\theta \int_0^{\infty} e^{-r^2} \cdot r \, dr = \pi/4,
 \end{aligned}$$

de donde resulta

$$\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

Es conveniente que el lector ubique en un diagrama los conjuntos O_1 y O_2 que corresponden en este ejemplo a los de igual denominación en el teorema (1.13).

En el espacio R^3 , las coordenadas de un punto $x = (x_1, x_2, x_3)$ se expresan en función de sus coordenadas polares (r, θ_1, θ_2) (figura 3) por medio de las fórmulas

$$(1.19) \quad \begin{cases} x_1 = r \cos \theta_1 \\ x_2 = r \sin \theta_1 \cos \theta_2 \\ x_3 = r \sin \theta_1 \sin \theta_2 \end{cases}$$

donde $r > 0$, $0 < \theta_1 < \pi$, $0 \leq \theta_2 < 2\pi$.

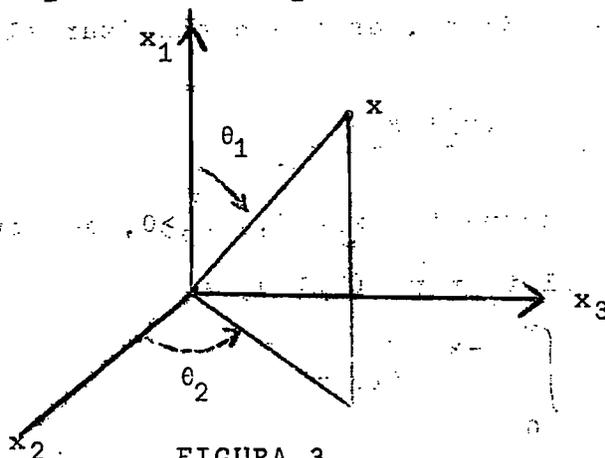


FIGURA 3

Nótese que, también en este caso, los números r , θ_1 y θ_2 están unívocamente determinados, dentro de sus respectivos intervalos, por medio de las fórmulas $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, $\cos \theta_1 = x_1/r$, $\cos \theta_2 = x_2/r \operatorname{sen} \theta_1$, $\operatorname{sen} \theta_2 = x_3 / r \operatorname{sen} \theta_1$, para lo cual es indispensable que se verifique $r \operatorname{sen} \theta_1 \neq 0$.

Las ecuaciones (1.19) definen una aplicación biyectiva del "prisma" $P = (0, \infty) \times (0, \pi) \times [0, 2\pi)$ que hemos representado en la figura 4 sobre el espacio R^3 menos la recta $x_2 = x_3 = 0$.

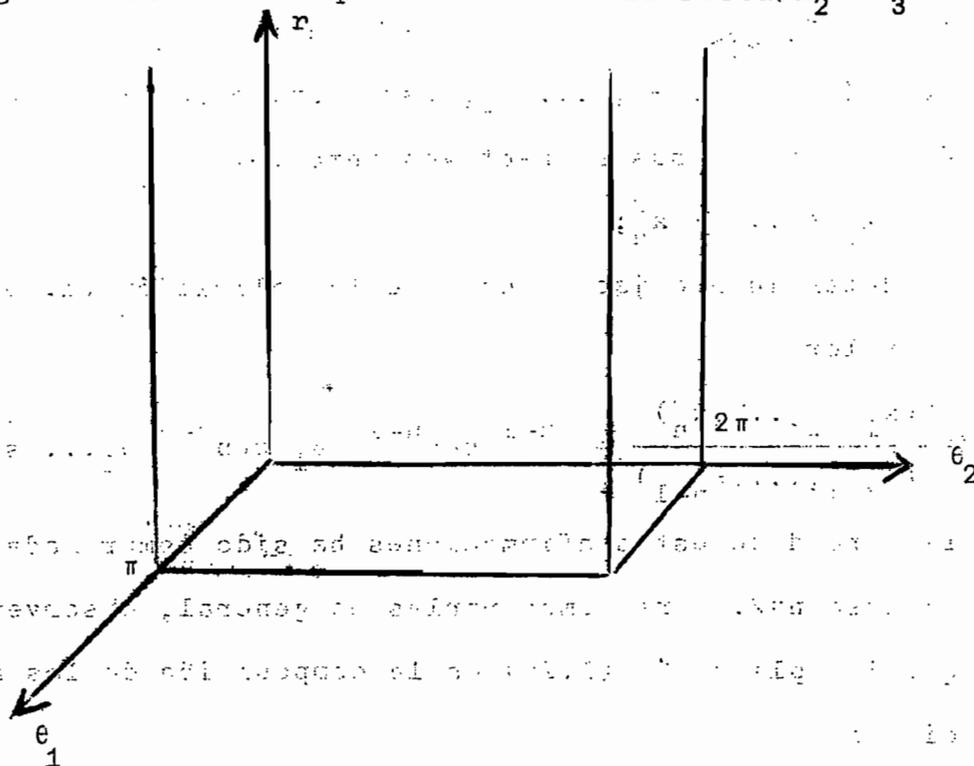


FIGURA 4

Los ejemplos anteriores nos inducen a considerar en el espacio R^n la aplicación definida por las ecuaciones

$$\begin{aligned}
 x_1 &= r \cos \theta_1 \\
 x_2 &= r \sin \theta_1 \cos \theta_2 \\
 x_3 &= r \sin \theta_1 \sin \theta_2 \cos \theta_3 \\
 &\dots \\
 x_{n-1} &= r \sin \theta_1 \dots \sin \theta_{n-2} \cos \theta_{n-1} \\
 x_n &= r \sin \theta_1 \dots \sin \theta_{n-2} \sin \theta_{n-1}
 \end{aligned}$$

$r > 0, 0 < \theta_1 < \pi, \dots, 0 < \theta_{n-2} < \pi, 0 \leq \theta_{n-1} < 2\pi.$

Por inducción sobre n demostraremos que:

- 1°) los números $r, \theta_1, \dots, \theta_{n-1}$ están unívocamente determinados dentro de sus respectivos intervalos;
- 2°) $r^2 = x_1^2 + \dots + x_n^2$;
- 3°) el determinante jacobiano J de la aplicación (1.20) tiene el valor

$$J = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(r, \theta_1, \dots, \theta_{n-1})} = r^{n-1} \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \dots \sin \theta_{n-2}$$

La verdad de estas afirmaciones ha sido comprobada en el caso $n=2$. Para demostrarlas en general, observemos que la aplicación (1.20) es la composición de las aplicaciones

$x_1 = x_1$	$x_1 = r \cos \theta_1$
$x_2 = r_1 \cos \theta_2$	$r_1 = r \sin \theta_1$
$x_3 = r_1 \sin \theta_2 \cos \theta_3$	$\theta_2 = \theta_2$
\dots	\dots
$x_n = r_1 \sin \theta_2 \dots \sin \theta_{n-1}$	$\theta_{n-1} = \theta_{n-1}$

de las cuales la primera es esencialmente la misma aplicación (1.20) en el espacio R^{n-1} ya que la coordenada x_1 queda fija, en tanto que la segunda representa una aplicación esencialmente idéntica a (1.18) que corresponde a la introducción de coordenadas polares en el plano:

Si suponemos que nuestras afirmaciones se verifican en el espacio R^{n-1} , entonces tendremos:

(a) las coordenadas del punto $(x_2, \dots, x_n) \in R^{n-1}$ determinan unívocamente a los números $r_1, \theta_2, \dots, \theta_{n-1}$ tales que $r_1 > 0$, $0 < \theta_2 < \pi, \dots, 0 < \theta_{n-2} < \pi, 0 < \theta_{n-1} < 2\pi$ y además, $r_1^2 = x_2^2 + \dots + x_n^2$;

(b) el punto (x_1, r_1) determina unívocamente a los números r y θ_1 que verifican $r > 0$ y $0 < \theta_1 < \pi$, pues $r_1 > 0$. Además, $r^2 = x_1^2 + r_1^2 = x_1^2 + x_2^2 + \dots + x_n^2$.

(c) en virtud de la regla de la cadena y la hipótesis inductiva,

$$\begin{aligned}
 J &= \frac{\partial(x_1, x_2, x_3, \dots, x_n)}{\partial(x_1, r_1, \theta_2, \dots, \theta_{n-1})} \cdot \frac{\partial(x_1, r_1, \theta_2, \dots, \theta_{n-1})}{\partial(r, \theta_1, \theta_2, \dots, \theta_{n-1})} \\
 &= \frac{\partial(x_2, x_3, \dots, x_n)}{\partial(r_1, \theta_2, \dots, \theta_{n-1})} \cdot \frac{\partial(x_1, r_1)}{\partial(r, \theta_1)} \\
 &= r_1^{n-2} \operatorname{sen}^{n-3} \theta_2 \dots \operatorname{sen} \theta_{n-2} \cdot r = r^{n-1} \operatorname{sen}^{n-2} \theta_1 \operatorname{sen}^{n-3} \theta_2 \dots \\
 &\quad \dots \operatorname{sen} \theta_{n-2}.
 \end{aligned}$$

Observación: Si se aceptan como posibles los valores extremos

0 y π para las coordenadas polares $r, \theta_1, \dots, \theta_{n-2}$, las fórmulas

(1.20) definen una aplicación (no inyectiva) cuya imagen es todo R^n .

La demostración de esta afirmación se realiza por el mismo camino inductivo.

Para obtener una aplicación inyectiva hemos restringido esta función más amplia al conjunto donde se verifica

$$r \operatorname{sen} \theta_1 \dots \operatorname{sen} \theta_{n-2} \neq 0,$$

lo que equivale a excluir del espacio R^n un conjunto que es unión finita de subespacios de dimensión $\leq n-2$, como se comprueba por una simple inspección a las ecuaciones (1.20).

Por simplicidad notacional es muy conveniente escribir la transformación (1.20) en la forma

$$x = r x',$$

donde r es la norma euclídeana o módulo del vector x , en tanto que

$$x' = (\cos \theta_1, \operatorname{sen} \theta_1 \cos \theta_2, \dots, \operatorname{sen} \theta_1 \dots \operatorname{sen} \theta_{n-1})$$

es un vector de módulo igual a uno.

El punto $\theta = (\theta_1, \dots, \theta_{n-1})$ al que llamaremos el argumento del vector x , varía en el intervalo $L \subset R^{n-1}$ definido por las relaciones

$$0 < \theta_1 < \pi, \dots, 0 < \theta_{n-2} < \pi, \quad 0 \leq \theta_{n-1} < 2\pi.$$

El determinante jacobiano J de la transformación (1.20) se puede escribir en la forma más concisa.

$$J = r^{n-1} g(\theta),$$

donde

$$g(\theta) = g(\theta_1, \dots, \theta_{n-1}) = \operatorname{sen}^{n-2} \theta_1 \operatorname{sen}^{n-3} \theta_2 \dots \operatorname{sen} \theta_{n-2}.$$

Denotando por R_+ el conjunto de los números reales no negativos, la integral de una función medible no negativa $f: R^n \rightarrow R$ se calcula entonces del modo siguiente:

$$(1.21) \quad \int_{R^n} f(x) dx = \int_{R_+} \int_L f(rx') r^{n-1} g(\theta) dr d\theta =$$

$$= \int_0^\infty dr r^{n-1} \int_L f(rx') g(\theta) d\theta.$$

La función $f(x)$ se llama una función radial si existe una función $f_0: R_+ \rightarrow R$, tal que $f(x) = f_0(|x|) = f_0(r)$.

En el caso particular de una función radial, la última fórmula nos da

$$(1.22) \quad \int_{R^n} f(x) dx = C_n \int_0^\infty f_0(r) r^{n-1} dr,$$

donde

$$C_n = \int_L g(\theta) d\theta.$$

Para calcular el valor de la constante C_n consideremos el caso de la función radial

$$f(x) = e^{-(x_1^2 + \dots + x_n^2)} = e^{-r^2}$$

cuya integral se calcula fácilmente por el teorema de Fubini:

$$\int_{R^n} f(x) dx = \prod_{k=1}^n \int_{-\infty}^{\infty} e^{-x_k^2} dx_k = C_n \int_0^\infty e^{-r^2} r^{n-1} dr =$$

$$= \frac{C_n}{2} \int_0^{\infty} t^{n/2-1} e^{-t} dt,$$

en virtud del cambio de variable $r = t^{1/2}$.

Recordando la definición de la función gama:

$$(1.23) \quad \Gamma(p) = \int_0^{\infty} t^{p-1} e^{-t} dt \quad (p > 0),$$

obtenemos

$$(1.24) \quad C_n = \frac{2 \pi^{n/2}}{\Gamma(n/2)}$$

En el tercer capítulo demostraremos que esta constante representa el área de la superficie esférica unitaria S^{n-1} , formada por todos los puntos x que verifican $|x| = 1$.

Para finalizar, haremos dos observaciones acerca de la función $\Gamma(p)$ definida por la integral (1.23). En primer lugar veamos por qué ella es finita si $p > 0$ y sólo en este caso.

Puesto que el integrando es positivo, tendremos

$$\int_0^{\infty} x^{p-1} e^{-x} dx = \int_0^1 + \int_1^{\infty} = I_1 + I_2.$$

La integral I_2 es finita cualquiera que sea el valor de p . En cuanto a I_1 , las relaciones obvias

$$e^{-1} x^{p-1} < x^{p-1} < e^{-x} x^{p-1} \quad (0 < x < 1),$$

muestran que I_1 es finita si y sólo si x^{p-1} es integrable sobre $(0,1)$ y esto ocurre si y sólo si $p > 0$.

La segunda observación, que puede saltarse sin perjuicio en una primera lectura de estas notas, está contenida en el siguiente lema:

(1.25) LEMA. Si a es un número real cualquiera, entonces para $p \rightarrow \infty$ se cumple

$$\frac{\Gamma(p)}{\Gamma(p+a)} \sim p^{-a}$$

en el sentido de que el cociente entre las dos expresiones separadas por el "ganchito" tiende a uno cuando $p \rightarrow \infty$.

Suponiendo primero que $a > 1$, en virtud del ejercicio 8 al final de este capítulo, tendremos

$$\begin{aligned} \frac{\Gamma(p) \Gamma(a)}{\Gamma(p+a)} &= \int_0^1 (1-x)^{a-1} x^{p-1} dx = \int_0^\infty ((1-e^{-t})^{a-1}) e^{-pt} dt = \\ &= \int_0^\infty t^{a-1} e^{-pt} dt - \int_0^\infty \{t^{a-1} - (1-e^{-t})^{a-1}\} e^{-pt} dt = \\ &= J_1 - J_2 . \end{aligned}$$

Por un lado, $J_1 = \Gamma(a) p^{-a}$. En cuanto a J_2 , las desigualdades

$$1 - e^{-t} < t \quad (t > 0), \quad 1 - e^{-t} > t - t^2/2 \quad (0 < t < 1)$$

muestran que el integrando de J_2 es positivo y además,

$$\begin{aligned} J_2 &= \int_0^\infty \{t^{a-1} - (1-e^{-t})^{a-1}\} e^{-pt} dt = \int_0^1 + \int_1^\infty \leq \\ &\leq \int_0^1 \{1 - (1-t/2)^{a-1}\} t^{a-1} e^{-pt} dt + \int_1^\infty t^{a-1} e^{-pt} dt < \end{aligned}$$

$$I_s < K \int_0^1 t^a e^{-pt} dt + \int_1^\infty t^a e^{-pt} dt < K \int_0^\infty t^a e^{-pt} dt = K \Gamma(a) p^{-a-1},$$

donde K es un número mayor que uno e independiente de p. Hemos usado el hecho de que la función continua

$$\frac{1-(1-t/2)^{a-1}}{t} \quad (0 < t \leq 1)$$

tiene un límite finito cuando $t \rightarrow 0$ y por consiguiente es acotada.

Para los demás valores de a, el lema sigue fácilmente de la fórmula $\Gamma(p+1) = p\Gamma(p)$ cuya demostración se pide en el ejercicio 9 al final del presente capítulo.

Ejercicios

1. Probar que $\Gamma(1/2) = \sqrt{\pi}$. Pista: introducir el cambio de variable $x = t^2$.

2. Demostrar la fórmula $\int_{|x| < a} f(x) dx = \int_0^a dr r^{n-1} \int_L f(rx') g(\theta) d\theta$.

3. Para qué valores reales de p es $|x|^p$ integrable sobre la bola unitaria $|x| < 1$?

4. Demuestre que si A es una matriz simétrica y $Q(x) = x A^t x$, donde ${}^t x$ denota el vector transpuesto de x , entonces la función $\exp(-Q(x))$ es integrable sobre R^n si y sólo si todos los autovalores de A son positivos; y en tal caso,

$$\int \exp(-Q(x)) dx = \frac{\pi^{n/2}}{\sqrt{\det A}}$$

Pista: introducir un cambio de variables $x = yU$, donde U es una matriz ortogonal, tal que UAU^{-1} es una matriz diagonal.

5. Demostrar que la integral biparamétrica

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$$

tiene un valor finito si $p > 0$ y $q > 0$ y verifica $B(p, q) = B(q, p)$. La función $B(p, q)$ se llama Función Beta.

6. Demostrar que la integral biparamétrica

$$\int_0^1 x^{p-1} |\log x|^{q-1} dx$$

es finita si $p > 0$ y $q > 0$ y expresar su valor en términos

de la función gama. Pista: introducir el cambio de variable $x = e^{-t}$.

7. Demostrar que la convolución de dos densidades es una densidad (la función medible $f: \mathbb{R}^n \rightarrow \mathbb{R}_+$ se llama una densidad si su integral sobre todo el espacio es igual a uno). Pista: usar el teorema de Fubini.

8. Para cada $p > 0$ considérese sobre \mathbb{R} la densidad

$$f_p(x) = \begin{cases} \frac{1}{\Gamma(p)} x^{p-1} e^{-x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

demostrar que si $p > 0$ y $q > 0$, entonces $f_p * f_q = f_{p+q}$ y obtener como "subproducto" la importantísima relación

$$B(p, q) = \frac{\Gamma(p) \Gamma(q)}{\Gamma(p+q)} \quad (p > 0, q > 0).$$

Pista:

$$(f_p * f_q)(x) = \int_{-\infty}^{\infty} f_p(x-y) f_q(y) dy = \int_0^x f_p(x-y) f_q(y) dy.$$

Reemplazar el integrando por su valor e introducir el cambio de variable $y = xt$.

9. Demostrar la relación $\Gamma(p+1) = p\Gamma(p)$. Pista: por el teorema de Beppo-Levi,

$$\Gamma(p+1) = \int_0^{\infty} x^p e^{-x} dx = \lim_{\varepsilon \rightarrow 0, A \rightarrow \infty} \int_{\varepsilon}^A x^p e^{-x} dx.$$

Integrar por partes la última integral y calcular el límite.

10. Probar las fórmulas: $(n+1)! = n!$ ($n=0,1,2,\dots$) y

$$\frac{(2k+1)!}{2^k} = \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{2^k} \sqrt{\pi} = \frac{(2k)!}{2^{2k} k!} \sqrt{\pi} \quad (k=1,2,\dots)$$

11. Expresar el valor de la integral $\int_0^1 \sqrt{1-x^3} dx$ en términos de la función gama. Pista: $x^3 = t$.

12. Expresar en términos de la función gama el valor de la integral

$$\int_0^{\infty} \frac{r^{n-1}}{(1+r^2)^{(n+1)/2}} dr \quad (n \text{ entero positivo})$$

Pista: introducir el cambio de variables $r = \sqrt{\frac{x}{1-x}}$ ($0 < x < 1$) y calcular con un poco de paciencia.

13. Usar el resultado del ejercicio precedente para calcular la integral sobre el espacio \mathbb{R}^n de la función radial

$$\frac{1}{(1 + |x|^2)^{(n+1)/2}} \quad (n \text{ entero positivo}).$$

14. Si $f(x)$ es una función integrable sobre \mathbb{R}^n y a un número real distinto de cero, $\int f(ax) dx = |a|^{-n} \int f(x) dx$.

15. Sea $A \subset \mathbb{R}^n$ un conjunto boreliano con la siguiente propiedad:

para cada $v \in \mathbb{R}^n$ con $|v|=1$, el conjunto $A_v = \{t \in \mathbb{R}^n : tv \in A\}$ es de medida nula. Probar que A es de medida nula.

Pista: usar la fórmula (1.21) y el teorema de Fubini para cambiar el orden de integración.

16. Si M es un conjunto convexo en \mathbb{R}^n , probar que la frontera de M tiene medida nula y que M es medible.

17. Introduciendo el cambio de variable $x = \sin^2 t$ en la integral del ejercicio 5, demostrar la fórmula:

$$\int_0^{\pi/2} \sin^{2p-1} t \cos^{2q-1} t dt = \frac{1}{2} B(p, q) \quad (p > 0, q > 0),$$

y en particular,

$$\int_0^{\pi/2} \sin^n x dx = \sqrt{\pi} \frac{\Gamma((n+1)/2)}{\Gamma((n+2)/2)} \quad (n=0, 1, 2, \dots)$$

18. La última relación del ejercicio precedente permite demostrar la interesante FORMULA DE WALLIS:

$$\frac{2}{\sqrt{\pi}} = \lim_{k \rightarrow \infty} \frac{(k!)^2 2^{2k}}{(2k)! \sqrt{k}}$$

Pista: expresar las integrales de las funciones $\sin^{2k-1} x$ y $\sin^{2k} x$ sobre el intervalo $(0, \pi)$, efectuar el cociente entre ambas y (hacer $k \rightarrow \infty$ usando) el lema (1.25) para demostrar que el cociente de las integrales tiende a uno.

19. Si $p: \mathbb{R}^n \rightarrow \mathbb{R}$ es un polinomio -o una función analítica- no constante, entonces para todo $c \in \mathbb{R}$, el conjunto $p^{-1}(c)$

tiene medida nula. (Sugerencia: ejercicio 15).

20. Sea $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal autoadjunta definida positiva. Calcular la medida del n -elipsoide

$$\{x: \langle Ax, Ax \rangle \leq 1\}.$$

CAPITULO II

LA INTEGRAL SOBRE UNA HIPERSUPERFICIE

1. Hipersuperficie

Si U es un conjunto abierto de \mathbb{R}^m y $\varphi: U \rightarrow \mathbb{R}^n$, escribiremos $\varphi = (\varphi_1, \dots, \varphi_n)$ y denotaremos por $u = (u_1, \dots, u_m)$ un punto cualquiera de U .

(2.1) DEFINICION. La aplicación inyectiva $\varphi: U \rightarrow \mathbb{R}^n$ se llama una inmersión si es de clase C^∞ y si la matriz jacobiana

$$\varphi'(u) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u_1} & \dots & \frac{\partial \varphi_1}{\partial u_m} \\ \dots & \dots & \dots \\ \frac{\partial \varphi_n}{\partial u_1} & \dots & \frac{\partial \varphi_n}{\partial u_m} \end{pmatrix}$$

tiene rango igual a m en cada punto de U .

En el caso particular en que se verifique $m=n-1$, escribiremos

$$J_k(U) = \frac{\partial(\varphi_1, \dots, \hat{\varphi}_k, \dots, \varphi_n)}{\partial(u_1, u_2, \dots, u_{n-1})} \quad (1 \leq k \leq n)$$

para indicar el determinante de la matriz cuadrada que se obtiene eliminando la k -ésima fila de $\varphi'(u)$.

Si $A \subset \mathbb{R}^n$ y $\varphi: A \rightarrow \mathbb{R}^n$, diremos que φ es de clase C^k si para cada $a \in A$ existe un entorno abierto U de a en \mathbb{R}^m y una aplicación $f: U \rightarrow \mathbb{R}^n$ de clase C^k , tal que $f|_{U \cap A} = \varphi|_{U \cap A}$.

(Ver A.3.7. para una definición equivalente).

(2.2) DEFINICION. Un conjunto no vacío $H \subset \mathbb{R}^n$, provisto de la topología inducida, se llama una hipersuperficie si para cada $p \in H$ existen conjuntos abiertos $S \subset H$ con $p \in S$ y $U \subset \mathbb{R}^{n-1}$ y una inmersión $\varphi: U \rightarrow \mathbb{R}^n$, tales que

$$1^\circ) \quad \varphi(U) = S$$

$$2^\circ) \quad \varphi: U \rightarrow S \text{ es un homeomorfismo.}$$

El par (U, φ) se denomina una carta local o sistema de coordenadas de H alrededor del punto p .

(2.3) TEOREMA. Un conjunto no vacío $H \subset \mathbb{R}^n$, provisto de la topología inducida, es una hipersuperficie si y sólo si para cada $p \in H$ existen conjuntos abiertos Ω y Ω^* de \mathbb{R}^n con $p \in \Omega$ y un difeomorfismo de clase C^∞ $f: \Omega \rightarrow \Omega^*$, tal que

$$\Omega \cap H = \{x \in \Omega : f_n(x) = 0\}$$

donde f_n es la última componente del vector $f = (f_1, \dots, f_n)$.

Demostración Sea H una hipersuperficie y sea (U, φ) una carta local alrededor del punto $p = \varphi(z)$ (figura 5)

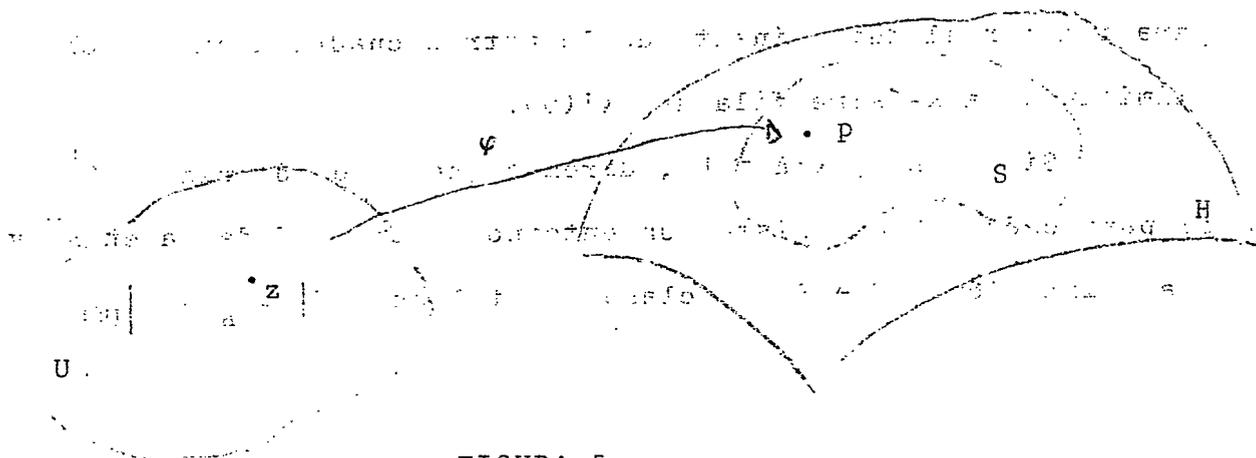


FIGURA 5

Puesto que $\psi'(z)$ tiene rango igual a $n-1$, existe un entero k , $1 \leq k \leq n$, tal que $J_k(z) \neq 0$. Consideremos la aplicación $g: U \times \mathbb{R} \rightarrow \mathbb{R}^n$ definida por

$$g(u, t) = \psi(u) + te_k,$$

donde $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ con el uno en la k -ésima coordenada. Luego,

$$g(u, 0) = \psi(u), \quad g(z, 0) = p,$$

$$\det g'(z, 0) = (-1)^{n+k} J_k(z) \neq 0.$$

Por el teorema de la aplicación inversa, existen un entorno abierto $U_1 \subset U$ del punto z , un número $\epsilon > 0$ y un entorno abierto A de p en \mathbb{R}^n , tales que

$$g: U_1 \times (-\epsilon, \epsilon) \rightarrow A$$

es un difeomorfismo de clase C^∞ con función inversa $f = g^{-1}$. Por ser ψ un homeomorfismo, $\psi(U_1)$ es un conjunto abierto de $S = \psi(U)$. Luego, existe un abierto B de \mathbb{R}^n , tal que $\psi(U_1) = B \cap H$. Si ahora ponemos $\Omega = A \cap B$, $\Omega^* = f(\Omega)$, se comprueba fácilmente que $p \in \Omega$ y

$$\Omega \cap H = A \cap B \cap H = \{x \in \Omega : f_n(x) = 0\}.$$

Escribiendo $\hat{f} = (f_1, \dots, f_{n-1})$ se tiene además que

$$\psi^{-1} \Big|_{\Omega \cap H} = \hat{f} \Big|_{\Omega \cap H}.$$

y en consecuencia $\psi^{-1}: S \rightarrow U$ es de clase C^∞ , pues una vez fijada la carta local (U, ψ) nuestros razonamientos se aplican a cualquier punto $p \in S$.

Recíprocamente, supongamos que para cada $p \in H$ existe un entorno abierto Ω de p en \mathbb{R}^n y un difeomorfismo $f: \Omega \rightarrow \Omega^*$ de clase C^∞ , tal que $\Omega \cap H = \{x \in \Omega : f_n(x) = 0\}$ (figura 6).

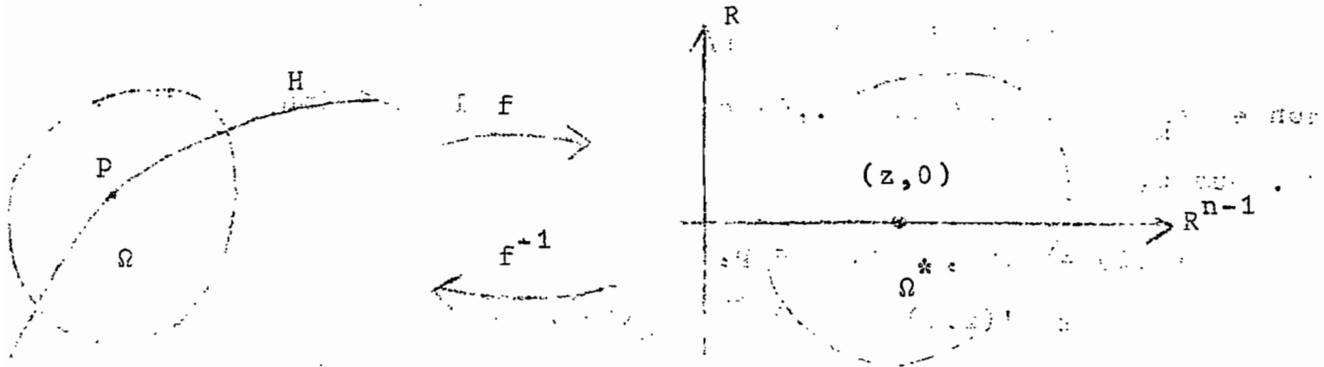


FIGURA 6

Si $U = \{u \in \mathbb{R}^{n-1} : (u, 0) \in \Omega^*\}$, entonces U es un conjunto abierto de \mathbb{R}^{n-1} y poniendo

$$\varphi(u) = f^{-1}(u, 0) \quad (u \in U)$$

obtenemos una carta local (U, φ) alrededor de p con $S = \varphi(U) = \Omega \cap H$. Que $\varphi'(u)$ tiene rango $n-1$ sigue fácilmente del hecho de que $Df^{-1}(u, 0)$ es invertible.

De los corolarios que siguen, el primero no hace más que destacar una conclusión obtenida en el curso de la demostración precedente.

(2.4) COROLARIO. Si (U, φ) es una carta local de una hipersuperficie H con $S = \varphi(U)$, entonces $\varphi^{-1}: S \rightarrow U$ es de clase C^∞ .

(2.5) COROLARIO. Sea H una hipersuperficie y (U_i, φ_i) ($i=1, 2$) dos cartas locales de H con $S = S_1 \cap S_2 \neq \emptyset$, donde $S_i = \varphi_i(U_i)$ (figura 7). Entonces los conjuntos $\varphi_i^{-1}(S)$ son abiertos en el

espacio R^{n-1} , y la aplicación

$$\psi_2 \circ \psi_1^{-1}: \psi_1^{-1}(S) \rightarrow \psi_2^{-1}(S)$$

es un difeomorfismo de clase C^∞

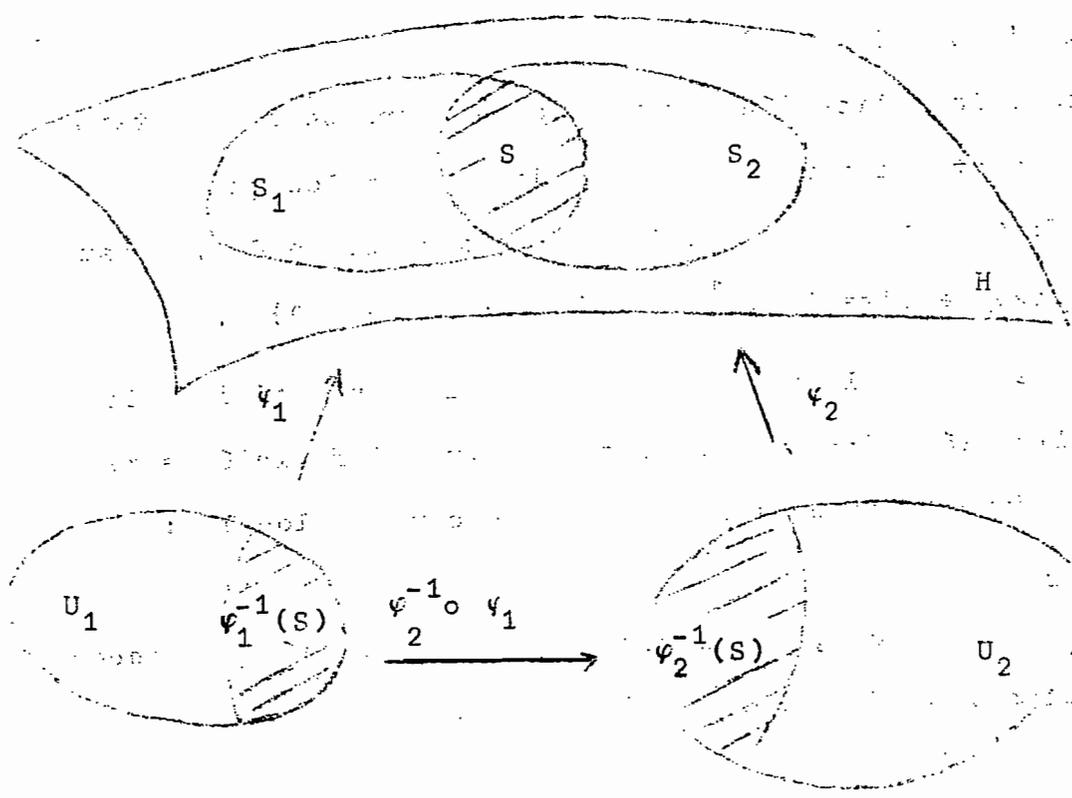


FIGURA 7

(2.6) PROPOSICION. Sea E un conjunto abierto de R^n y sea $\phi: E \rightarrow R$ una aplicación de clase C^∞ que verifica $\text{grad } \phi \neq 0$ en cada punto de E . Si $0 \in \phi(E)$, entonces $\phi^{-1}(0)$ es una hipersuperficie.

Antes de dar la demostración, aclaremos que

$$\text{grad } \phi = (D_1 \phi, \dots, D_n \phi) = \left(\frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n} \right).$$

Si $p \in \phi^{-1}(0)$ podemos suponer sin pérdida de generalidad que $D_n \phi(p) \neq 0$.

Entonces la aplicación $f: E \rightarrow \mathbb{R}^n$ definida por

$$f(x) = (x_1, \dots, x_{n-1}, \phi(x))$$

verifica

$$\det f'(p) = D_n \phi(p) \neq 0,$$

y el teorema de la aplicación inversa nos asegura de la existencia de dos conjuntos abiertos Ω y Ω^* de \mathbb{R}^n , tales que $p \in \Omega \subset E$ y $f: \Omega \rightarrow \Omega^*$ es un difeomorfismo de clase C^∞ . Puesto que $f_n(x) = \phi(x)$, resulta $\Omega \cap \phi^{-1}(0) = \{x \in \Omega: f_n(x) = 0\}$.

Señalemos finalmente que para una hipersuperficie H la noción de aplicación $f: H \rightarrow \mathbb{R}$ de clase C^k (ver definición antes de 2.2) puede formularse en términos de las cartas locales:

Proposición 2.7

Sea H una hipersuperficie, y sea $f: H \rightarrow \mathbb{R}$ una aplicación. Entonces son equivalentes las afirmaciones siguientes:

- f es de clase C^k
- Para toda carta local (U, ψ) de H la aplicación $f \circ \psi: U \rightarrow \mathbb{R}$ es de clase C^k .

Demostración: Es muy fácil ver que la composición de aplicaciones de clase C^k es nuevamente una aplicación de clase C^k ; por lo tanto, a) \Rightarrow b) puesto que cada $\psi: U \rightarrow \mathbb{R}^n$ es de clase C^∞ .

Para probar la recíproca, sea $p \in H$; veremos que hay un abierto

Ω en \mathbb{R}^n tal que $p \in \Omega$ y una $\psi: \Omega \rightarrow \mathbb{R}$ de clase C^k tal que

$$f|_{\Omega \cap H} = \psi|_{\Omega \cap H}$$

Ciertamente hay una carta (U, φ) de H con $p \in U$; por la primera parte de la demostración de 2.3 se sabe que hay un abierto $U_1 \subset U$ en \mathbb{R}^{n-1} , un $\varepsilon > 0$ y un entorno abierto Ω de p en \mathbb{R}^n tales que:

Sea $g: U_1 \times (-\varepsilon, \varepsilon) \rightarrow \Omega$ un difeomorfismo de clase C^∞ , con $g(u, t) = \varphi(u) + t e_k$; la aplicación inversa (de clase C^∞) será de la forma

$\omega \rightarrow (h_1(\omega), h_2(\omega))$ donde $h_1: \Omega \rightarrow U_1$ y $h_2: \Omega \rightarrow (-\varepsilon, \varepsilon)$ son de clase C^∞ .

Ponemos $\psi: \Omega \rightarrow \mathbb{R}$ por $\psi(\omega) = f \circ h_1(\omega)$; como $\omega \in \Omega \cap H$ equivale a $h_2(\omega) = 0$, resulta que $\psi \circ \varphi^{-1} = f \circ \varphi^{-1}$ y por lo tanto ψ coincide con f sobre $\varphi(U_1) = \Omega \cap H$.

Por otro lado como $f \circ \varphi^{-1}: U_1 \rightarrow \mathbb{R}$ es de clase C^k (por hipótesis) y h_1 es de clase C^∞ es claro que $\psi: \Omega \rightarrow \mathbb{R}$ es de clase C^k .

La propiedad b) de 2.7 suele adoptarse como definición de aplicación de clase C^k en el contexto general de la teoría de variedades diferenciables.

2. Espacio tangente a una hipersuperficie

Sea H una hipersuperficie de \mathbb{R}^n y sea (U, φ) una carta local alrededor del punto $p = \varphi(z)$. Puesto que φ es una inmersión, el subespacio $D\varphi(z) \subset (\mathbb{R}^{n-1})$ es un hiperplano de \mathbb{R}^n al que denotaremos por H_p . Si (V, ψ) es otra carta local alrededor del mismo punto $p = \psi(\bar{z})$, podemos suponer que $\varphi(U) = \psi(V)$.

Aplicando la regla de la cadena a la función $\varphi = \psi \circ (\psi^{-1} \circ \varphi)$

en el punto $z \in U$ resulta

$$D\varphi(z) = D\psi(\psi^{-1} \circ \varphi(z)) \circ D(\psi^{-1} \circ \varphi)(z) = D\psi(\bar{z}) \circ D(\psi^{-1} \circ \varphi)(z).$$

Teniendo en cuenta que $\psi^{-1} \circ \varphi: U \rightarrow V$ es un difeomorfismo de clase C^∞ , concluimos que $D(\psi^{-1} \circ \varphi)(z)$ es un isomorfismo del espacio vectorial \mathbb{R}^{n-1} y por consiguiente,

$$H_p = D\varphi(z)(\mathbb{R}^{n-1}) = D\psi(\bar{z})(\mathbb{R}^{n-1}).$$

Es decir, el subespacio H_p es independiente de la carta local que consideremos alrededor del punto p .

(2.7) DEFINICION. El subespacio H_p se denomina el espacio tangente a H en el punto p . Los vectores $v \in H_p$ se llaman vectores tangentes a H en p .

Notemos que si $\{e_1, \dots, e_{n-1}\}$ es la base canónica de \mathbb{R}^{n-1} , los vectores

$$V_k(u) = D\varphi(u)(e_k) \quad (1 \leq k \leq n-1)$$

forman una base del espacio tangente $H_p(u)$.

Recordando que la matriz de la aplicación lineal

$D\varphi(u)$ es $\varphi'(u) = (\partial\varphi_i / \partial u_k)$, si denotamos por $\{E_1, \dots, E_n\}$ la base canónica de \mathbb{R}^n , tendremos

$$V_k(u) = \sum_{i=1}^n \frac{\partial \varphi_i}{\partial u_k} E_i = \left(\frac{\partial \varphi_1}{\partial u_k}, \dots, \frac{\partial \varphi_n}{\partial u_k} \right) = \frac{\partial \varphi}{\partial u_k}.$$

El conjunto $p + H_p$ (figura 8) se llama el hiperplano afín tangente a H en el punto p .

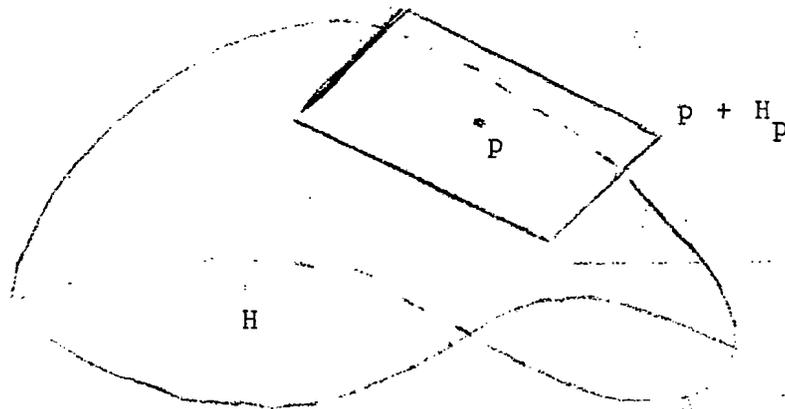


FIGURA 8

Si (U, φ) es una carta local alrededor del punto $p = \varphi(u)$, denotaremos por $N(u)$ un vector unitario ortogonal al subespacio $H_{\varphi(u)}$ y por $g(u)$ el volumen (i.e., la medida) del paralelepípedo generado por los vectores $V_1(u), \dots, V_{n-1}(u), N(u)$ (figura 9).

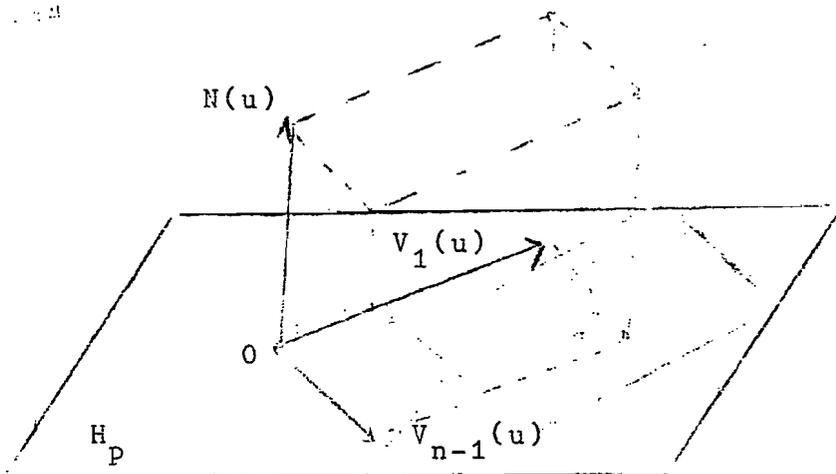


FIGURA 9

De acuerdo con los resultados del primer capítulo, tendremos

$$(2.8) \quad g(u) = \left| \det \begin{pmatrix} N(u) \\ v_1(u) \\ \vdots \\ v_{n-1}(u) \end{pmatrix} \right|^2$$

La función $g:U \rightarrow \mathbb{R}_+$ desempeña un papel fundamental en la definición del elemento de área del sector $S = \varphi(U)$. Nos proponemos demostrar que denotando por $\langle \cdot, \cdot \rangle$ el producto escalar entre vectores de \mathbb{R}^n , su valor se calcula por medio de las fórmulas

$$(2.9) \quad g(u) = \sqrt{\sum_{k=1}^n J_k^2(u)}$$

$$(2.10) \quad g(u) = \sqrt{\det (g_{ik}(u))}$$

donde

$$J_k(u) = \frac{\partial(\psi_1, \dots, \psi_k, \dots, \psi_n)}{\partial(u_1, u_2, \dots, u_{n-1})}$$

$$g_{ik}(u) = \langle V_i(u), V_k(u) \rangle.$$

Para demostrar (2.9) consideremos la función lineal $f_u: \mathbb{R}^n \rightarrow \mathbb{R}$ definida por

$$f_u(x) = \det \begin{pmatrix} x \\ V_1(u) \\ \vdots \\ V_{n-1}(u) \end{pmatrix} \quad (x \in \mathbb{R}^n).$$

Luego, $f_u(E_k) = (-1)^{k+1} J_k(u)$ y si definimos el vector $\omega(u)$ por medio de la fórmula

$$\omega(u) = (f_u(E_1), \dots, f_u(E_n)),$$

para cada $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ tendremos

$$\begin{aligned} f_u(x) &= f_u\left(\sum_{k=1}^n x_k E_k\right) = \sum_{k=1}^n x_k f_u(E_k) = \\ &= \langle x, \omega(u) \rangle. \end{aligned}$$

Puesto que $\omega(u)$ es ortogonal a cada uno de los vectores $V_1(u), \dots, V_{n-1}(u)$, existe un número real $\lambda(u)$ tal que $\omega(u) = \lambda(u) N(u)$ y en virtud de (2.8),

$$\begin{aligned} g(u) &= |f_u(N(u))| = |\langle N(u), \omega(u) \rangle| = \\ &= |\lambda(u)| = |\omega(u)| = \sqrt{\sum_{k=1}^n J_k^2(u)}. \end{aligned}$$

Para demostrar (2.10) consideremos en el espacio $H_{\psi}(u)$ una base ortonormal $\omega_1(u), \dots, \omega_{n-1}(u)$ y pongamos

$$V_i(u) = \sum_{k=1}^{n-1} \alpha_{ik}(u) \omega_k(u).$$

Denotando por ${}^t\alpha(u)$ la matriz transpuesta de la matriz $\alpha(u) = (\alpha_{ik}(u))$, se tienen, las relaciones $(g_{ik}(u)) = \alpha(u) {}^t\alpha(u)$,

$$\begin{bmatrix} N(u) \\ V_1(u) \\ \vdots \\ V_{n-1}(u) \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha(u) \end{bmatrix} \begin{bmatrix} N(u) \\ \omega_1(u) \\ \vdots \\ \omega_{n-1}(u) \end{bmatrix}$$

Tomando determinantes en estas relaciones y teniendo en cuenta que la última matriz es ortogonal se obtiene (2.10).

3. Elemento de área de una hipersuperficie

Sean (U, φ) y (V, ψ) dos cartas locales de una hipersuperficie H y supongamos por brevedad en la notación que $\varphi(U) = \psi(V)$. De no ser así tomaríamos la intersección de estos conjuntos y en lugar de U y V las imágenes inversas de dicha intersección por las aplicaciones φ y ψ , respectivamente, tal como en el corolario (2.5).

Si para cada $v = (v_1, \dots, v_{n-1}) \in V$ definimos en forma análoga

$$\bar{V}_k(v) = D\psi(v)(e_k), \quad \bar{g}_{ik}(v) = \langle \bar{V}_i(v), \bar{V}_k(v) \rangle,$$

$$\bar{g}(v) = \sqrt{\det(\bar{g}_{ik}(v))},$$

por ser $h = \psi^{-1} \circ \varphi : U \rightarrow V$ un difeomorfismo y observando que

$$(g_{ik}(u)) = {}^t\varphi'(u) \varphi'(u), \quad (\bar{g}_{ik}(v)) = {}^t\psi'(v) \psi'(v),$$

si aplicamos la regla de la cadena de la función $\varphi = \psi \circ h$,

obtendremos muy fácilmente

$$(g_{ik}(u)) = {}^t h'(u) (\bar{g}_{ik}(h(u))) h'(u)$$

de donde resulta

$$(2.11) \quad g(u) = \bar{g}(h(u)) |\det h'(u)| = \bar{g}(h(u)) \left| \frac{\partial v}{\partial u} \right|;$$

y en virtud de la fórmula del cambio de variables,

$$(2.12) \quad \int_V \bar{g}(v) dv = \int_U g(u) du.$$

(2.13) DEFINICION. El valor común de las integrales (2.12) se llama el área del sector $S = \varphi(U)$; y el elemento de área $d\sigma$ del sector S se define simbólicamente por la expresión

$$(2.14) \quad d\sigma = g(u) du$$

4. Integral sobre una hipersuperficie

Si H es una hipersuperficie y $f: H \rightarrow \mathbb{R}$, diremos que f es medible si para cualquier carta local (U, φ) de H , la composición $f \circ \varphi$ es medible sobre U (comparar con 2.7.b)). Un conjunto $E \subset H$ se llama medible si su función característica χ_E es medible. El soporte $\text{sup } f$ de una función $f: H \rightarrow \mathbb{R}$ se define como la clausura (en H) del conjunto de todos los puntos $x \in H$, tales que $f(x) \neq 0$.

Sean (U, φ) y (V, ψ) dos cartas locales de H con $S_1 = \varphi(U)$ y $S_2 = \psi(V)$. Si $f: H \rightarrow \mathbb{R}$ es una función medible no negativa cuyo soporte está contenido en $S = S_1 \cap S_2$, la fórmula del cambio de variables y la fórmula (2.11) muestran que

$$\int_U f \circ \varphi(u) g(u) du = \int_{\varphi^{-1}(S)} f \circ \varphi(u) g(u) du =$$

$$= \int_{\psi^{-1}(S)} f \circ \psi(v) \bar{g}(v) dv = \int_V f \circ \psi(v) \bar{g}(v) dv.$$

El valor común de estas integrales es, por definición, la integral de f sobre H y se escribe

$$(2.15) \quad \int_H f d\sigma = \int_U f \circ \psi(u) g(u) du,$$

siempre que $\text{sop } f \subset \psi(U)$.

A fin de extender la noción de integral sobre H a cualquier función medible no negativa $f: H \rightarrow \mathbb{R}$ cuyo soporte no está contenido en ninguna carta local, es fundamental el siguiente lema cuya demostración daremos en un apéndice al final de estas notas:

(2.16) LEMA. Si $(S_i, i \in I)$ es un cubrimiento abierto de H , entonces existe una familia de funciones continuas $h_i: H \rightarrow \mathbb{R}$ ($i \in I$), tales que:

- el soporte de h_i está contenido en S_i ;
- cada punto de H posee un entorno cuya intersección con $\text{sop } h_i$ es vacío para todos los índices i con la posible excepción de un número finito de ellos;
- $h_i \geq 0$ y $\sum h_i = 1$.

La familia de funciones $(h_i, i \in I)$ se llama una partición continua de la unidad subordinada al cubrimiento abierto $(S_i, i \in I)$ (ver A.3.4., A.3.5.)

Puesto que H es un espacio métrico separable y los sistemas locales de coordenadas alrededor de cada punto forman un cubrimiento abierto de H , existe una familia numerable de cartas locales (U_i, φ_i) , $i \in I$, tal que los correspondientes sectores $S_i = \varphi_i(U_i)$ forman un cubrimiento abierto de H . Si $(h_i, i \in I)$ es una partición continua de la unidad subordinada a este cubrimiento, para cualquier función medible no negativa $f: H \rightarrow \mathbb{R}$ definimos la integral de f sobre H por medio de la fórmula

$$(2.17) \quad \int_H f \, d\sigma = \sum_i \int_H f h_i \, d\sigma,$$

donde cada término se evalúa según la fórmula (2.15) por vía de la carta local (U_i, φ_i) , pues $\text{supp}(f h_i) \subset S_i$.

Lo primero es verificar que la definición es correcta, es decir, independiente del cubrimiento y de la partición de la unidad a él subordinada. En efecto, si (V_j, ψ_j) , $j \in J$, es otra familia numerable de cartas locales cuyos correspondientes sectores $\psi_j(V_j)$ forman un cubrimiento de H y si $(k_j, j \in J)$ es una partición de la unidad subordinada a este nuevo cubrimiento, tendremos

$$\sum_i \int_H f h_i \, d\sigma = \sum_i \int_H f h_i \left(\sum_j k_j \right) d\sigma = \sum_i \sum_j \int_H f h_i k_j \, d\sigma$$

y análogamente,

$$\sum_j \int_H f k_j \, d\sigma = \sum_j \sum_i \int_H f k_j h_i \, d\sigma;$$

pero en las series dobles de términos no negativos el orden de las sumatorias es irrelevante, con lo cual queda demostrado que la

definición (2.17) es correcta.

El área $|E|$ de un conjunto medible $E \subset H$ se define como la integral de su función característica:

$$|E| = \int_H \chi_E \, d\sigma.$$

Diremos que la función medible $f: H \rightarrow \mathbb{R}$ es integrable sobre H si las funciones no negativas $f^+ = \sup(0, f)$ y $f^- = \sup(0, -f)$ tienen integrales finitas sobre H y en tal caso definimos la integral de f sobre H por medio de la fórmula.

$$\int_H f \, d\sigma = \int_H f^+ \, d\sigma - \int_H f^- \, d\sigma.$$

Como en el estudio de la integral de Lebesgue en \mathbb{R}^n , las fórmulas integrales valen en general una vez demostradas para funciones no negativas.

5. Expresiones particulares del elemento de área

A veces es posible escribir las ecuaciones que definen una carta local (U, φ) en la forma

$$(2.18) \quad \varphi(u_1, \dots, u_{n-1}) = (u_1, \dots, u_{n-1}, \psi(u_1, \dots, u_{n-1})),$$

o, lo que es lo mismo, en la forma

$$x_1 = u_1, \dots, x_{n-1} = u_{n-1}, \quad x_n = \psi(u_1, \dots, u_{n-1}),$$

donde el punto $u = (u_1, \dots, u_{n-1})$ varía en la proyección ortogonal del sector $S = \varphi(U)$ sobre el hiperplano $x_n = 0$. En tal caso, ponien-

do

$$\psi_k = D_k \psi = \frac{\partial \psi}{\partial u_k} \quad (k=1, 2, \dots, n-1),$$

según un sencillo cálculo, el elemento de área de S es

$$(2.19) \quad d\sigma = \sqrt{\psi_1^2 + \dots + \psi_{n-1}^2 + 1} \, du.$$

Para un sector S dado por las ecuaciones (2.18), cuyos puntos satisfacen además una relación de la forma

$$\phi(x) = \phi(x_1, \dots, x_n) = c,$$

donde c es una constante y ϕ una función de clase C^∞ que verifica $\phi_n = D_n \phi \neq 0$ en cada punto de S, tendremos

$$\phi(u_1, \dots, u_{n-1}, \psi(u_1, \dots, u_{n-1})) = c,$$

de donde, derivando con respecto a u_k , resulta $\phi_k + \phi_n \psi_k = 0$, es decir,

$$\psi_k = - \frac{\phi_k}{\phi_n}$$

y por consiguiente,

$$(2.20) \quad d\sigma = \frac{\sqrt{\phi_1^2 + \dots + \phi_{n-1}^2}}{|\phi_n|} \, du = \frac{|\text{grad } \phi|}{|\phi_n|} \, du,$$

donde todas las funciones que figuran deben evaluarse en el punto $(u, \psi(u))$. La expresión obtenida es de fundamental importancia para demostrar el lema que estudiaremos en la siguiente sección.

6. Resolución de integrales múltiples

El propósito de esta sección es demostrar el siguiente lema

(2.21) LEMA. Sea E un conjunto abierto de \mathbb{R}^n y sea $\phi: E \rightarrow \mathbb{R}$ una función de clase C^∞ que verifica $\text{grad } \phi \neq 0$ en cada punto de E . Si $f(x) = f(x_1, \dots, x_n)$ es una función medible no negativa sobre dicho conjunto, entonces

$$\int_E f(x) dx = \int_{\phi(E)} dt \int_{\phi^{-1}(t)} \frac{f}{|\text{grad } \phi|} d\sigma_t,$$

donde $d\sigma_t$ representa el elemento de área de la hipersuperficie $\phi^{-1}(t)$.

Por hipótesis, en cada punto de E es distinta de cero, al menos una de las derivadas parciales $\phi_k = D_k \phi$.

Supongamos que en cierto punto $p \in E$ se verifica $\phi_n(p) \neq 0$. Entonces la aplicación $(x_1, \dots, x_n) \rightarrow (u_1, \dots, u_{n-1}, t)$ dada por

$$(2.22) \quad u_1 = x_1, \dots, u_{n-1} = x_{n-1}, t = \phi(x_1, \dots, x_n),$$

cuyo determinante jacobiano

$$\frac{\partial(u_1, \dots, u_{n-1}, t)}{\partial(x_1, x_2, \dots, x_n)} = \phi_n$$

es distinto de cero en el punto p , establece un difeomorfismo entre cierto entorno abierto $V \subset E$ de dicho punto y cierto conjunto abierto $W \subset \mathbb{R}^n$. La aplicación inversa de (2.22) se puede escribir en la forma

$$(2.23) \quad x_1 = u_1, \dots, x_{n-1} = u_{n-1}, x_n = \psi(u_1, \dots, u_{n-1}, t).$$

Con el fin de fijar ideas, supongamos por un momento que

$V = E$ y pongamos

$$W_t = \{u \in \mathbb{R}^{n-1} : (u, t) \in W\}.$$

Si $u \in W_t$, entonces el punto $x = (u, \psi(u, t))$ verifica

$$\phi(x) = \phi(u, \psi(u, t)) = t.$$

Por consiguiente, para cada t fijo, las ecuaciones (2.23) proveen un representación paramétrica o carta local de hipersuperficie $\phi^{-1}(t)$. Se entiende que para ello, el punto u ha de variar en W_t .

En virtud de la fórmula del cambio de variables y de la expresión (2.20),

$$\begin{aligned} \int_E f(x) dx &= \int_W f(u, \psi(u, t)) \frac{1}{|\phi_n|} du dt = \\ &= \int dt \int_{W_t} \frac{f(u, \psi(u, t))}{|\text{grad } \phi|} \frac{|\text{grad } \phi|}{|\phi_n|} du = \\ &= \int_{\phi(E)} dt \int_{\phi^{-1}(t)} \frac{f}{|\text{grad } \phi|} d\sigma_t. \end{aligned}$$

Si $V \neq E$, la fórmula recién obtenida se mantiene válida a condición de que el soporte de f esté contenido en V . En efecto, en tal caso,

$$\begin{aligned} \int_E f(x) dx &= \int_V f(x) dx = \int_{\phi(V)} dt \int_{V \cap \phi^{-1}(t)} \frac{f}{|\text{grad } \phi|} d\sigma_t = \\ &= \int_{\phi(E)} dt \int_{\phi^{-1}(t)} \frac{f}{|\text{grad } \phi|} d\sigma_t, \end{aligned}$$

pues $t \in \phi(V)$ si y sólo si $V \cap \phi^{-1}(t) \neq \emptyset$ y f es nula fuera de V .

En el caso general podremos considerar un cubrimiento abierto numerable (V_i) del conjunto E con la propiedad de que la fórmula del lema es válida para cualquier función $f \geq 0$ cuyo soporte esté contenido en algún V_i . Si (h_i) es una partición continua de la unidad subordinada a dicho cubrimiento, entonces para cualquier función medible $f \geq 0$ tendremos

$$\begin{aligned} \int_E f(x) dx &= \int_E f(\sum_i h_i) dx = \sum_i \int_E f h_i dx = \\ &= \sum_i \int_{\phi(E)} dt \int_{\phi^{-1}(t)} \frac{f h_i}{|\text{grad } \phi|} d\sigma_t = \\ &= \int_{\phi(E)} dt \int_{\phi^{-1}(t)} \frac{f}{|\text{grad } \phi|} d\sigma_t. \quad \text{Q.E.D.} \end{aligned}$$

La fórmula demostrada representada, en cierto sentido, una generalización del teorema de Fubini, pues si elegimos $E = \mathbb{R}^n$ y $\phi(x) = x_n$, entonces $\phi^{-1}(t)$ es el hiperplano $x_n = t$, la expresión (2.19) del elemento de área en este caso particular nos da $d\sigma_t = du$, y la fórmula del lema toma la forma

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}} dt \int_{\mathbb{R}^{n-1}} f(u, t) du.$$

La generalización consiste en la posibilidad de substituir dichos hiperplanos por hipersuperficie más generales.

Ejercicios

1. Sea H una hipersuperficie y $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una isometría, es decir, una aplicación de la forma $Tx = xA + b$, donde A es una matriz ortogonal. Demostrar que si $d\sigma_T$ es el elemento de área de la hipersuperficie TH , entonces para cada función medible no negativa $f: TH \rightarrow \mathbb{R}$ se cumple

$$\int_{TH} f d\sigma_T = \int_H f \circ T d\sigma.$$

En particular, si $E \subset H$ es medible, entonces $|TE| = |E|$.

Pista. usar (2.10) para demostrar que $d\sigma_T = d\sigma$, teniendo en cuenta que si (U, φ) es una carta local de H , entonces $(U, T \circ \varphi)$ es una carta local de TH y observar que es suficiente demostrar la fórmula para una función cuyo soporte está contenido en una de estas cartas locales.

Nota. La propiedad enunciada en este ejercicio se expresa comúnmente diciendo que el elemento de área es invariante bajo isometrías.

2. Probar que la superficie esférica unitaria $S^{n-1} = \{x \in \mathbb{R}^n : |x|=1\}$ es una hipersuperficie.

Pista: aplicar la proposición (2.6) a la función

$$\phi(x) = x_1^2 + \dots + x_n^2 - 1 \text{ sobre el conjunto } E = \{x: x \neq 0\}.$$

3. Mostrar que si $f(x)$ es una función integrable sobre S^{n-1} y A una matriz ortogonal, entonces

$$\int_{S^{n-1}} f(xA) d\sigma = \int_{S^{n-1}} f d\sigma.$$

The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that proper record-keeping is essential for the integrity of the financial system and for the ability to detect and prevent fraud.

Internal Controls



Internal controls are designed to ensure that the organization's resources are used efficiently and effectively. They provide a framework for the organization to achieve its objectives and to minimize the risk of error or fraud. Management is responsible for designing and implementing internal controls, while the board of directors is responsible for overseeing the process.

The board of directors has a key role to play in the internal control process. It should establish the overall control environment and should be involved in the design and implementation of internal controls. The board should also monitor the effectiveness of internal controls and should take corrective action where necessary.

Management is responsible for the day-to-day operation of internal controls. It should ensure that internal controls are designed to address the organization's risks and that they are implemented effectively. Management should also monitor the effectiveness of internal controls and should take corrective action where necessary.

The board of directors should also be involved in the monitoring of internal controls. It should receive regular reports from management on the effectiveness of internal controls and should take corrective action where necessary. The board should also be involved in the design and implementation of internal controls.

CAPITULO III

APLICACIONES

1. El elemento de área de la hipersuperficie S^{n-1}

En el primer capítulo hemos visto que las ecuaciones que definen a las coordenadas polares en R^n se pueden escribir en la forma

$$(3.1) \quad x = rx'$$

donde r es el módulo del vector x , en tanto que el punto

$$x' = (\cos \theta_1, \text{sen } \theta_1 \cos \theta_2, \dots, \text{sen } \theta_1 \dots \text{sen } \theta_{n-1})$$

varía sobre la superficie esférica unitaria S^{n-1} formada por todos los puntos de R^n de módulo igual a uno. Nótese que el punto x' es una función explícita del argumento $\theta = (\theta_1, \dots, \theta_{n-1})$, el cual haremos variar, en este capítulo, en el intervalo abierto $LC R^{n-1}$ definido por las desigualdades

$$0 < \theta_1 < \pi, \dots, 0 < \theta_{n-2} < \pi, \quad 0 < \theta_{n-1} < 2\pi.$$

En estas condiciones, la aplicación $\theta \rightarrow x'$ de L en R^n provee una parametrización o carta local de S^{n-1} que según veremos enseguida abarca a todos los puntos de esta hipersuperficie con la excepción de un conjunto de área nula. En el caso de R^3 , los puntos excluidos de S^2 forman un arco de meridiano.

Si M es un subintervalo cualquiera de L , definido por ciertas relaciones

$$\alpha_1 < \theta_1 < \beta_1, \dots, \alpha_{n-1} < \theta_{n-1} < \beta_{n-1},$$

la medida de Lebesgue del conjunto

$$\omega = \{x \in \mathbb{R}^n : r < R, \theta \in M\} \quad (\text{figura 10})$$

se calcula muy fácilmente usando el jacobiano de la transformación (3.1), a saber:

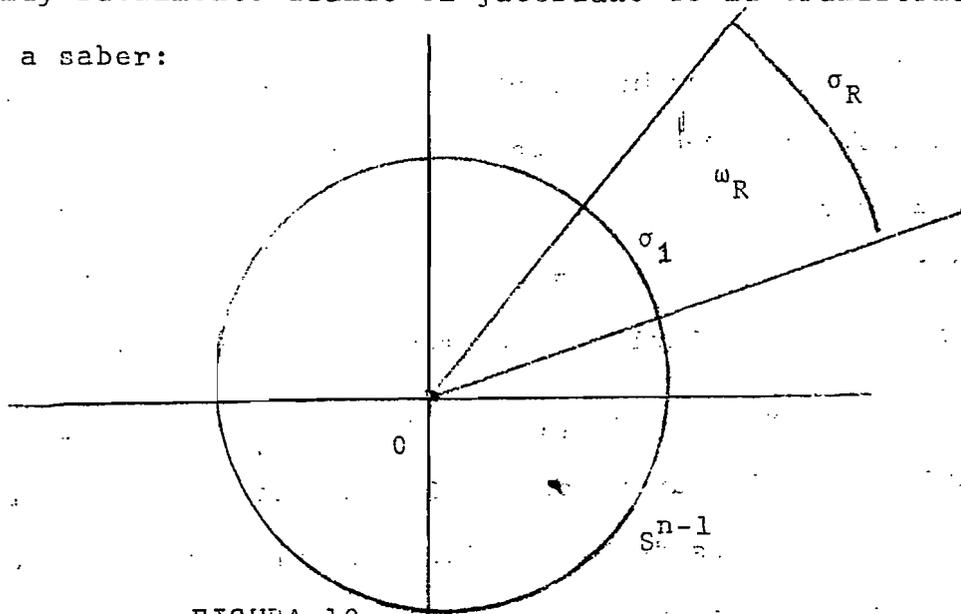


FIGURA 10

$$(3.2) \quad |\omega_R| = \int \omega_R dx = \int_0^R dr r^{n-1} \int_M g(\theta) d\theta,$$

donde

$$(3.3) \quad g(\theta) = \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \dots \sin \theta_{n-2}$$

Para calcular el elemento de área de la hipersuperficie σ_R dada por las ecuaciones paramétricas

$$(3.4) \quad x = Rx' \quad (\theta \in M)$$

debemos formar los determinantes

$$J_i = \frac{\partial(x_1, \dots, \hat{x}_i, \dots, x_n)}{\partial(\theta_1, \theta_2, \dots, \theta_{n-1})} = R^{n-1} T_i \quad (i=1, 2, \dots, n),$$

donde la variable con la cuspide se omite y $T_i = T_i(\theta)$ representa

un polinomio trigonométrico cuya expresión exacta no necesitamos. Luego,

$$d\sigma_R = \left(\sum_{i=1}^n J_i^2 \right)^{1/2} d\theta = R^{n-1} h(\theta) d\theta,$$

donde $h = h(\theta)$ es una función continua sobre L . Por consiguiente, el área del sector σ_R es

$$|\sigma_R| = R^{n-1} \int_M h(\theta) d\theta = R^{n-1} |\sigma_1|.$$

Por otro lado, si aplicamos el lema (2.21) con $E = \omega_R$, $f =$ función característica de ω_R y $\phi(x) = |x| = r$, teniendo en cuenta que en este caso $|\text{grad } \phi| = 1$, resulta

$$|\omega_R| = \int_{\omega_R} 1 \cdot dx = \int_0^R dt \int_{\sigma_t} d\sigma_t = \int_0^R |\sigma_t| dt.$$

Luego,

$$(3.5) \quad \frac{d|\omega_R|}{dR} = |\sigma_R| = R^{n-1} \int_M h(\theta) d\theta.$$

Pero de (3.2) se deduce que también

$$\frac{d|\omega_R|}{dR} = R^{n-1} \int_M g(\theta) d\theta.$$

Esto muestra que las funciones continuas g y h tienen la misma integral sobre cualquier subintervalo M de L , de donde $g = h$. Por consiguiente, la expresión

$$d\sigma_1 = g(\theta) d\theta$$

representa el elemento de área de la hipersuperficie S^{n-1} , al menos en lo que respecta a la carta local $\theta \rightarrow x'$ dada por las ecuaciones

$$\begin{aligned}
 (3.6) \quad x'_1 &= \cos \theta_1 \\
 x'_2 &= \text{sen } \theta_1 \cos \theta_2 \\
 &\dots \dots \dots \\
 x'_{n-1} &= \text{sen } \theta_1 \dots \text{sen } \theta_{n-2} \cos \theta_{n-1} \\
 x'_n &= \text{sen } \theta_1 \dots \text{sen } \theta_{n-2} \text{sen } \theta_{n-1} \quad (0 \in L).
 \end{aligned}$$

De (3.5) se deduce inmediatamente la relación

$$(3.7) \quad |\omega_R| = \frac{R^n}{|n!|} |\sigma_1|.$$

Quedan por aclarar dos importantes detalles que hemos dejado para el final con el objeto de presentar rápidamente la idea fundamental:

1º) de la relación (3.7) se deduce que cada conjunto de área nula en S^{n-1} genera un cono con vértice en el origen de medida también nula y recíprocamente, si el cono generado por un conjunto $E \subset S^{n-1}$ tiene medida nula, entonces E es un conjunto de área nula. Si en las ecuaciones (3.6) se hace que alguno de los ángulos θ_i valga 0 ó π se obtiene un subconjunto de S^{n-1} que es la intersección de esta hipersuperficie con una unión finita de subespacios de dimensión $\leq n-1$, es decir, un conjunto de medida nula en R^n . Esto demuestra que la carta local (3.6) abarca toda la hipersuperficie S^{n-1} con la excepción de un conjunto de área nula (téngase presente que área y medida son invariantes bajo rotaciones del espacio);

2º) la aplicación (3.4) representa una inmersión de L en R^n . En efecto, si desarrollamos por la primera columna el determinante jacobiano de la aplicación (3.1) con $r = R$,

tendremos

$$R^{n-1}g(\theta) = \frac{\partial(x_1, x_2, \dots, x_n)}{\partial(R, \theta_1, \dots, \theta_{n-1})} = \sum_{i=1}^n (-1)^{1+i} J_i.$$

Puesto que g no se anula en ningún punto de L , si $R \neq 0$ entonces alguno de los J_i debe ser distinto de cero, lo cual demuestra que la matriz Jacobiana de la aplicación (3.4) tiene rango igual a $n-1$. Por consiguiente, (3.6) es realmente una carta local de S^{n-1} , tal como habíamos afirmado al comienzo.

La primera consecuencia de todo lo anterior es que el área total de la hiper-superficie S^{n-1} es igual a la constante C_n de la fórmula (1.24), es decir,

$$(3.8) \quad |S^{n-1}| = \frac{2 \pi^{n/2}}{\Gamma(n/2)}$$

De acuerdo con (3.7), la medida (o volumen) V_n de la bola unitaria de R^n es el número

$$(3.9) \quad V_n = (1/n) |S^{n-1}| = \frac{\pi^{n/2}}{\Gamma((n+2)/2)},$$

y la medida de la bola de radio r se obtiene sencillamente multiplicando este número por r^n .

Para dar mayor coherencia a la notación usual en ciertas fórmulas de uso muy frecuente, vamos en adelante a denotar por $dx' = g(\theta)d\theta$ el elemento de área de la superficie esférica unitaria.

Con esta notación, la fórmula (1.21) para calcular la integral de una función medible no negativa $f: R^n \rightarrow R$ se puede escribir

en la siguiente forma que trae consigo una interesante interpretación geométrica:

$$(3.10) \quad \int_{\mathbb{R}^n} f(x) dx = \int_0^\infty dr r^{n-1} \int_L f(rx') g(\theta) d\theta =$$

$$= \int_0^\infty dr r^{n-1} \int_{S^{n-1}} f(rx') dx'.$$

Tal vez sin mayores consecuencias pero igualmente curioso es el hecho de que el volumen V_n de la bola unitaria tiende a cero cuando $n \rightarrow \infty$. Para demostrarlo conviene distinguir dos casos, según que n sea par o impar.

La sencilla idea que hemos usado para demostrar la fórmula $d\sigma_1 = g(\theta) d\theta$ ha sido expuesta en [2].

2. Densidades de probabilidad

Una función medible no negativa $f: \mathbb{R}^n \rightarrow \mathbb{R}$ se llama una densidad de probabilidad o más brevemente una densidad si su integral sobre (todo el) espacio es igual a uno. La densidad f está concentrada en el conjunto $A \subset \mathbb{R}^n$ si es nula fuera de dicho conjunto.

Recordemos que un vector $X = (X_1, \dots, X_n)$ cuyas componentes son variables aleatorias sobre cierto espacio de probabilidad (Ω, \mathcal{A}, P) se llama un vector aleatorio. Un vector aleatorio puede definirse como una función $X: \Omega \rightarrow \mathbb{R}^n$ con la propiedad

de que para cualquier conjunto boreliano E de \mathbb{R}^n la imagen

inversa $X^{-1}(E)$ es un miembro de la σ -álgebra \mathcal{G} .

Se dice que el vector aleatorio X está distribuido según la densidad f y se escribe $X \sim f$ si para cada conjunto boreliano E de \mathbb{R}^n se verifica

$$P\{X \in E\} = P(X^{-1}(E)) = \int_E f(x) dx.$$

Nos limitaremos a considerar vectores aleatorios que admiten una densidad de probabilidad por ser los únicos que conciernen al tema que estamos estudiando. Cada vez que se tomen dos variables aleatorias X e Y , se suponer que el vector (X, Y) tiene una densidad de probabilidad $f(x, y)$; y análogamente para más de dos variables. En este contexto las demostraciones representan sencillos ejercicios de cambios de variables que ilustran sobre la utilización del método.

A pesar de dicha limitación, hemos procurado que el desarrollo teórico sea coherente con las definiciones adoptadas.

Supongamos que X está distribuido según la densidad f concentrada en el conjunto abierto A y sea $\varphi: A \rightarrow B$ un difeomorfismo de clase C^1 sobre cierto conjunto abierto B de \mathbb{R}^n . La cuestión que se plantea naturalmente en esta situación es: ¿cuál es la densidad del vector aleatorio $Y = \varphi(X)$?

La respuesta es inmediata sobre la base del teorema (1.13). En primer lugar es claro que la densidad de Y está concentrada en B . Además, para cada conjunto boreliano $E \subset B$, tendremos

$$P\{Y \in E\} = P\{X \in \psi^{-1}(E)\} = \int_{\psi^{-1}(E)} f(x) dx =$$

$$= \int_E f(\psi^{-1}(y)) \left| \frac{\partial x}{\partial y} \right| dy,$$

donde $\partial x / \partial y$ representa el determinante jacobiano de la aplicación $x = \psi^{-1}(y)$. Por consiguiente, el vector aleatorio

Y está distribuido según la densidad g dada por

$$(3.11) \quad g(y) = \begin{cases} f(\psi^{-1}(y)) \left| \frac{\partial x}{\partial y} \right| & (y \in B) \\ 0 & (y \notin B). \end{cases}$$

Por ejemplo, si la variable aleatoria X tiene una densidad f concentrada en el intervalo $(0, \infty)$, entonces $Y = \sqrt{X}$ tiene la densidad $2y f(y^2)$ concentrada en el mismo intervalo.

Algunas veces la aplicación $\psi: A \rightarrow B$ no es inyectiva pero es posible expresar el conjunto A como unión numerable de ciertos conjuntos A_j , abiertos y disjuntos dos a dos, de modo que para cada j , la restricción

$$\psi_j = \psi|_{A_j}$$

es un difeomorfismo de clase C^1 de A_j sobre B . Por ejemplo, si

$\psi: \mathbb{R} \rightarrow \mathbb{R}$ es la función $\psi(x) = x^2$, podemos elegir $A_1 = (-\infty, 0)$,

$A_2 = (0, \infty)$, $B = (0, \infty)$. En dichas condiciones, si E es un subconjunto boreliano de B , entonces tendremos

$$\begin{aligned}
 P\{Y \in E\} &= P\{X \in \varphi^{-1}(E)\} = \\
 &= \int_{\varphi^{-1}(E)} f(x) dx = \sum_j \int_{\varphi_j^{-1}(E)} f(x) dx = \\
 &= \sum_j \int_E f(\varphi_j^{-1}(y)) \left| \frac{\partial \varphi_j^{-1}}{\partial y} \right| dy,
 \end{aligned}$$

donde $\partial \varphi_j^{-1} / \partial y$ representa el determinante jacobiano de la aplicación φ_j^{-1} evaluado en el punto $y \in E$.

Permutando suma e integral, lo que es posible en virtud del teorema de Beppo-Levi, obtenemos como densidad del vector Y la función

$$(3.12) \quad g(y) = \sum_j f(\varphi_j^{-1}(y)) \left| \frac{\partial \varphi_j^{-1}}{\partial y} \right| \quad (y \in B).$$

Volviendo al ejemplo de la función $\varphi(x) = x^2$ ($x \in \mathbb{R}$), para la variable aleatoria $Y = X^2$ se obtiene fácilmente la densidad

$$(3.13) \quad \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} \quad (y > 0)$$

Otro ejemplo interesante es el de la variable $Y = \sin X$, cuya densidad está concentrada en el intervalo $B = (-1, 1)$. En este caso conviene elegir, como A_j , el intervalo $((2j-1)\pi/2, (2j+1)\pi/2)$, $j = 0, \pm 1, \pm 2, \dots$ y como φ_j la aplicación $\sin_j: A_j \rightarrow B$ igual a la restricción de la función $\varphi(x) = \sin x$ al intervalo A_j , con lo que se obtiene para la densidad de Y la expresión

$$g(y) = (1-y^2)^{-1/2} \sum_{j=-\infty}^{\infty} f(\text{sen}_j^{-1} y) \quad (|y| < 1).$$

Si f es la densidad del vector $X = (X_1, \dots, X_n)$, la función g definida por la fórmula

$$(3.14) \quad g(x_1, \dots, x_m) = \int_{R^{n-m}} f(x_1, \dots, x_n) dx_{m+1} \dots dx_n$$

representa la densidad del vector $Y = (X_1, \dots, X_m)$.

Las densidades de la forma (3.14), obtenidas a partir de f por integración parcial de esta función, se llaman densidades marginales.

Las variables aleatorias X_i ($i=1, 2, \dots, n$) se llaman (mutuamente) independientes si la densidad f del vector $X = (X_1, \dots, X_n)$ es de la forma

$$f(x) = f_1(x_1) f_2(x_2) \dots f_n(x_n),$$

donde f_i representa la densidad de X_i .

Ejemplo 1. Sea X una variable aleatoria. Si $X \sim f$ e $Y = aX + b$ con $a \neq 0$, entonces

$$(3.15) \quad Y \sim \frac{1}{|a|} f\left(\frac{y-b}{a}\right)$$

Ejemplo 2. Para hallar la densidad de la suma $S = X + Y$ de dos variables aleatorias, supongamos que $(X, Y) \sim f(x, y)$ y consideremos la aplicación

$$\psi: \begin{cases} s = x + y \\ t = y \end{cases}$$

cuya inversa es

$$\varphi^{-1} : \begin{cases} x = s-t \\ y = t. \end{cases}$$

En virtud de (3.11), la densidad del vector $(S, T) = \varphi(X, Y) = (X + Y, Y)$ es $f(s-t, t)$ pues el jacobiano de φ^{-1} es idéntico a uno. Por lo tanto,

$$(3.16) \quad S \sim \int_{-\infty}^{\infty} f(s-t, t) dt,$$

En el caso particular, en que X e Y sean independientes, tendremos $f(x, y) = g(x) \cdot h(y)$, de donde

$$S \sim \int_{-\infty}^{\infty} g(s-t) \cdot h(t) dt ;$$

es decir, la densidad de la suma de dos variables aleatorias independientes es la convolución de las correspondientes densidades.

La densidad de un vector aleatorio puede modificarse arbitrariamente dentro de un conjunto de medida nula sin dejar de ser una densidad para el mismo vector. De esta observación vamos a sacar partido en el siguiente.

Ejemplo 3: Para hallar la densidad del producto $U = XY$ de dos variables aleatorias, supongamos que $(X, Y) \sim f(x, y)$ y que f es nula en los puntos situados sobre los ejes coordenados, o sea que f está concentrada en $\{(x, y) : x \neq 0, y \neq 0\}$. Consideremos ahora la aplicación

$$\varphi : \begin{cases} u = xy \\ v = y \end{cases} \quad (x \neq 0, y \neq 0)$$

cuya inversa es

$$\varphi^{-1} : \begin{cases} x = u/v \\ y = v \end{cases} \quad (u \neq 0, v \neq 0)$$

En virtud de (3.11), el vector $(U, V) = \varphi(X, Y) = (XY, Y)$ tiene la densidad

de donde resulta
$$\frac{1}{|v|} f\left(\frac{u}{v}, v\right)$$

$$(3.17) \quad U \sim \int_{-\infty}^{\infty} \frac{1}{|v|} f\left(\frac{u}{v}, v\right) dv.$$

Por lo dicho, antes, la hipótesis acerca de f no representa ninguna restricción.

Ejemplo 4: Hallar la densidad del cociente $T = X/Y$ de dos variables aleatorias independientes, sabiendo que la densidad de Y está concentrada en la semirrecta $(0, \infty)$.

Suponiendo que $X \sim f$ e $Y \sim g$, tendremos

$$(X, Y) \sim f(x) g(y),$$

donde $g(y) = 0$ si $y < 0$.

Consideremos la aplicación

$$\varphi : \begin{cases} t = x/y \\ v = y \end{cases} \quad (y > 0)$$

cuya inversa es

$$\varphi^{-1}: \begin{cases} x = tv \\ y = v \end{cases} \quad (v > 0).$$

Entonces el vector $(T, V) = \varphi(X, Y) = (X/Y, Y)$ tiene la densidad

$$f(tv) g(v)v \quad (v > 0),$$

de donde resulta

$$(3.18) \quad T \sim \int_0^{\infty} f(tv) g(v) v \, dv.$$

El valor medio o promedio EX de una variable aleatoria X con densidad f es la integral

$$(3.19) \quad EX = \int_{-\infty}^{\infty} xf(x) \, dx$$

siempre que esta integral exista en el sentido de Lebesgue, es decir, siempre que $x f(x)$ sea absolutamente integrable sobre \mathbb{R} .

Si existe el valor medio $\mu = EX$, la varianza $\text{Var } X$ se define por la fórmula

$$(3.20) \quad \text{Var } X = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) \, dx$$

Supongamos que $(X, Y) \sim f(x, y)$ y que existen los valores medios

$$EX = \int_{-\infty}^{\infty} dx \, x \int_{-\infty}^{\infty} f(x, y) \, dy, \quad EY = \int_{-\infty}^{\infty} dy \, y \int_{-\infty}^{\infty} f(x, y) \, dx.$$

Por consiguiente,

$$EX + EY = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f(x, y) \, dx \, dy =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} sf(s-t, t) ds dt =$$

$$= \int_{-\infty}^{\infty} ds \int_{-\infty}^{\infty} f(s-t, t) dt = E(X+Y),$$

en virtud del cambio de variable

$$s = x+y, \quad t = y$$

y de la fórmula (3.16).

La fórmula (3.15) permite probar la relación $E(aX) = a EX$ siempre que exista el valor medio EX y que el número a sea distinto de cero. Para librarnos de esta última restricción es suficiente ampliar la definición de valor medio poniendo $E(a) = a$ para la variable aleatoria que toma el valor constante a en cada punto de Ω (las variables aleatorias constantes no tienen densidad, por lo cual su valor medio requiere una definición aparte). Nótese que $E(X+b) = EX + E = EX + E1$, nuevamente en virtud de (3.15).

Sobre la base de (3.17) se demuestra análogamente que si X e Y son variables aleatorias independientes y existen los valores medios EX y EY , entonces $EX \cdot EY = E(XY)$. En efecto, en este caso, la densidad del vector (X, Y) es de la forma $f(x, y) = g(x) h(y)$, de donde

$$\begin{aligned} EX \cdot EY &= \int_{-\infty}^{\infty} xg(x) dx \int_{-\infty}^{\infty} yh(y) dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u f\left(\frac{u}{v}, v\right) \frac{1}{|v|} du dv \end{aligned}$$

en virtud del cambio de variables

$$u = xy, \quad v = y \quad (x \neq 0, y \neq 0).$$

Por consiguiente,

$$EX \cdot EY = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} \frac{1}{|v|} f\left(\frac{u}{v}, v\right) dv = E(XY).$$

Muy fácilmente se demuestran las fórmulas:

$$(3.21) \quad \text{Var}(aX + b) = a^2 \text{Var} X;$$

si $\mu = EX$, $\nu = EY$, entonces

$$(3.22) \quad \text{Var}(X+Y) = \text{Var} X + \text{Var} Y +$$

$$+ 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu)(y-\nu) f(x,y) dx dy.$$

Si X e Y son independientes, la integral doble del último término es nula y nos queda $\text{Var}(X + Y) = \text{Var} X + \text{Var} Y$.

Muchas veces es más fácil calcular la función de distribución

$$F(x) = P\{X < x\} = \int_{-\infty}^x f(t) dt$$

de la variable aleatoria X que tratar de calcular directamente su densidad f . Una vez calculada F , la densidad se obtiene inmediatamente por diferenciación: $f(x) = F'(x)$ en casi todo punto. Por ejemplo, si f es la densidad y F la función de distribución de la variable aleatoria X , e $Y = X^2$, entonces, para cada $y > 0$,

$$G(y) = P\{Y < y\} = P\{|X| < \sqrt{y}\} = P\{-\sqrt{y} < X < \sqrt{y}\} = F(\sqrt{y}) - F(-\sqrt{y})$$

y por derivación resulta nuevamente como densidad de Y la función

$$\frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} \quad (y > 0)$$

A partir de esta expresión se demuestran fácilmente las siguientes afirmaciones:

- 1°) si $EX = 0$, entonces $EX^2 = \text{Var } X$;
- 2°) si existe el valor medio $\mu = EX$, entonces $\text{Var } X = \text{Var}(X - \mu) = E(X - \mu)^2$.

La covarianza de dos variables aleatorias X e Y con valores medios μ y ν , respectivamente, se define por la fórmula

$$\text{Cov}(X, Y) = E[(X - \mu)(Y - \nu)]$$

siempre que el segundo miembro exista. En particular, $\text{Cov}(X, X) = \text{Var } X$.

EJERCICIOS

- 1) Demostrar por derivación que si X tiene una densidad f concentrada en $(0, \infty)$, entonces $Y = \sqrt{X}$ tiene la densidad $2y f(y^2)$ ($y > 0$).
- 2) Demostrar la fórmula $\text{Var}(X + Y) = \text{Var } X + \text{Var } Y + 2 \text{Cov}(X, Y)$.
- 3) Probar que si $(X, Y) \sim f(x, y)$, $\mu = EX$, $\nu = EY$, entonces

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu) (y-\nu) f(x, y) dx dy$$

4) Aplicando la desigualdad de Cauchy-Schwartz, probar que

$$|\text{Cov}(X, Y)| \leq \sqrt{\text{Var } X} \cdot \sqrt{\text{Var } Y}$$

Pista: $f(x, y) = \sqrt{f(x, y)} \cdot \sqrt{f(x, y)}$.

3. Algunas densidades útiles en Estadística

La densidad unidimensional más importante del cálculo de probabilidades es la densidad gaussiana:

$$(3.23) \quad f(x) = \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} x^2 \right]$$

Si $X \sim f$, entonces $Y = \sigma X + m$ ($\sigma > 0$) tiene la densidad

$$(3.24) \quad \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{(x-m)^2}{2\sigma^2} \right]$$

que recibe el nombre de densidad normal con promedio m y varian-za σ^2 . y se denota por $N(m, \sigma^2)$. En particular, $N(0, 1)$ es la densidad (3.23).

Es fácil demostrar que si $X \sim N(0, 1)$, entonces $EX=0$, $\text{Var } X = 1$ y por consiguiente $Y = \sigma X + m$ verifica $EY = m$, $\text{Var } Y = \sigma^2$.

Vamos a calcular dos densidades de probabilidad que desempeñan un papel importantísimo en Estadística.

(a) Sean X_1, \dots, X_n variables aleatorias independientes con la misma densidad normal $N(0,1)$. La variable aleatoria

$$X^2 = X_1^2 + \dots + X_n^2$$

se llama la variable "ji cuadrado" con n grados de libertad.

Para obtener su función de densidad, notemos que por la

hipótesis de independencia, el vector $X = (X_1, \dots, X_n)$ tiene la

densidad

$$f(x) = \frac{1}{(\sqrt{2\pi})^n} \exp \left[-\frac{x_1^2 + \dots + x_n^2}{2} \right] =$$

$$= \frac{1}{(2\pi)^{n/2}} \exp \left[-\frac{r^2}{2} \right]$$

donde $r^2 = x_1^2 + \dots + x_n^2$. Por consiguiente, para cada $u > 0$ tenemos

$$\begin{aligned} P\{X^2 < u\} &= \int_{|x| < \sqrt{u}} f(x) dx = \\ &= \frac{C_n}{(2\pi)^{n/2}} \int_0^{\sqrt{u}} (r^{n-1} \exp[-\frac{r^2}{2}]) dr = \\ &= \frac{C_n}{2(2\pi)^{n/2}} \int_0^u t^{\frac{n}{2}-1} \exp[-\frac{t}{2}] dt, \end{aligned}$$

donde C_n es la constante (1,24).

Por derivación resulta que la densidad de X^2 es la función

$$(3.25) \quad \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} u^{\frac{n}{2}-1} \exp[-\frac{u}{2}] \quad (u > 0).$$

Se entiende que esta densidad está concentrada en el

semieje positivo. De (3.25) se sigue que la densidad de la variable aleatoria

$$Y = \sqrt{\chi^2/n}$$

es la función

$$(3.26) \quad \frac{2^{-n/2}}{2^{n/2} \Gamma(\frac{n}{2})} y^{n-1} \exp \left[-\frac{ny^2}{2} \right] \quad (y > 0).$$

b) Supongamos ahora que X e Y son variables aleatorias independientes, la primera con densidad normal N(0,1) y la segunda con la densidad (3.26). Para calcular la densidad del cociente T = X/Y, la fórmula (3.18) nos da

$$c \int_0^\infty \exp \left[-\frac{t^2 v^2}{2} \right] v^{n-1} \exp \left[-\frac{nv^2}{2} \right] v \, dv = \dots$$

$$= c \int_0^\infty v^n \exp \left[-\frac{v^2}{2} (n+t^2) \right] dv,$$

donde

$$c = \frac{1}{\sqrt{2\pi}} \frac{2^{-n/2}}{2^{n/2} \Gamma(n/2)}$$

Introduciendo en la última integral la nueva variable

$$u = \frac{v}{2} (n + t^2) = \frac{n v^2}{2} \left(1 + \frac{t^2}{n} \right) \quad (t \text{ fijo}),$$

resulta que la densidad de T está dada por la función

$$(3.27) \quad \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} \left(1 + \frac{t^2}{n} \right)^{-(n+1)/2}$$

que es la célebre densidad "t de Student" con n grados de libertad.

Para futuras referencias conviene que denotemos por $\chi^2(n)$ la densidad (3.24) y por $t(n)$ la densidad (3.27). Esta notación pone en evidencia el único parámetro del que dependen esas dos familias de densidades, a saber, el número n de grados de libertad.

En la siguiente sección resultarán muy útiles los siguientes lemas:

(3.28) LEMA. Si las componentes del vector $X = (X_1, X_2, \dots, X_n)$ son mutuamente independientes, lo mismo es cierto para el vector $(X_1 + X_2, X_3, \dots, X_n)$.

(3.29) LEMA. La suma de dos variables aleatorias independientes con densidades normales tiene densidad normal.

El primero de estos lemas se demuestra calculando la densidad del vector $(Y_1, Y_2, \dots, Y_n) = (X_1 + X_2, X_2, \dots, X_n)$ con ayuda de (3.11) y luego la densidad marginal correspondiente al vector (Y_1, Y_3, \dots, Y_n) .

El siguiente lema afirma que la convolución de dos densidades normales $N(m_k, \sigma_k^2)$, $k = 1, 2$, es una densidad normal; y es suficiente probarlo en el caso $m_1 = m_2 = 0$, pues la familia de funciones $N(m, \sigma^2)$, $m \in \mathbb{R}$, $\sigma^2 > 0$, es invariante bajo traslaciones (también bajo cambios de escala).

La transformada de Fourier o función característica de una densidad f sobre \mathbb{R} es la función

$$\hat{f}(t) = (Ff)(t) = \int_{-\infty}^{\infty} \exp[itx] f(x) dx ;$$

y en el caso de la densidad normal,

$$F(N(0, \sigma^2)) = \exp \left[-\frac{1}{2} \sigma^2 t^2 \right].$$

Por consiguiente, si $f_k = N(0, \sigma_k^2)$, $k=1,2$, entonces tendremos

$$\begin{aligned} F(f_1 * f_2) &= \exp \left[-\frac{1}{2} \sigma_1^2 t^2 \right] \exp \left[-\frac{1}{2} \sigma_2^2 t^2 \right] = \\ &= \exp \left[-\frac{1}{2} (\sigma_1^2 + \sigma_2^2) t^2 \right] = F(N(0, \sigma_1^2 + \sigma_2^2)). \end{aligned}$$

EJERCICIO

Demstrar que si las componentes del vector (X_1, \dots, X_n) son independientes, lo mismo es cierto para el vector (X_1^2, X_2, \dots, X_n) .

Pista: considerar la aplicación $\psi: \mathbb{R}^n \rightarrow \mathbb{R}_+ \times \mathbb{R}^{n-1}$ definida por $\psi(x) = (x_1^2, x_2, \dots, x_n)$ y usar (3.12) con $A_1 = \{x: x_1 > 0\}$, $A_2 = \{x: x_1 < 0\}$ y $\psi_j = \psi|_{A_j}$.

Observación: Aplicando inductivamente el resultado de este ejercicio y el lema (3.28), resulta que si las componentes del vector (X_1, \dots, X_n) son independientes, lo mismo es cierto para los vectores

$(X_1 + \dots + X_k, X_{k+1}, \dots, X_n)$ y $(X_1^2, \dots, X_k^2, X_{k+1}, \dots, X_n)$, $k = 2, 3, \dots, n-1$.

4. La densidad normal en \mathbb{R}^n y el teorema de R. Fischer

El valor medio de una matriz rectangular $X = (X_{ij})$ cuyas componentes son variables aleatorias se define por la fórmula

$$EX = (EX_{ij})$$

siempre que existan los valores medios de cada una de ellas. Obviamente,

$$E(X+Y) = EX + EY, \quad E(AX) = A(EX) \quad \text{y}$$

$$E(XB) = (EX)B$$

siempre que las matrices constantes A y B, así como la matriz aleatoria Y posean las dimensiones necesarias como para que las operaciones indicadas tengan sentido.

La matriz de covarianzas del vector aleatorio $X = (X_1, \dots, X_n)$ con valor medio $EX = 0$ es la matriz de n filas por n columnas:

$$\text{Cov } X = E(X X^T)$$

y en el caso general definiremos $\text{Cov } X = \text{Cov}(X-EX)$; de modo que en la posición (i,j) de dicha matriz figura el valor $\text{Cov}(X_i, X_j)$ y en la diagonal principal las varianzas $\text{Var } X_i$ de las variables que componen el vector X.

Es obvio que $\text{Cov } X$ es una matriz simétrica. Demostraremos que además es positiva definida, a menos que el vector X esté concentrado en una variedad lineal de dimensión $< n$, en cuyo caso X no admite ninguna densidad y queda fuera de nuestra consideración. En efecto, si $L = a_1 X_1 + \dots + a_n X_n$ es una

combinación lineal de las variables X_i , a las cuales por brevedad supondremos con promedio nulo, entonces

$$E(L^2) = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j) a_i a_j$$

y es claro que la anulación de esta expresión implica que $L = 0$ con probabilidad uno, lo que significa que X está concentrado en el subespacio $a_1 x_1 + \dots + a_n x_n = 0$.

Si A es una matriz constante de n filas por n columnas, $Cov(XA) = {}^t A (Cov X) A$.

Diremos que el vector $X = (X_1, \dots, X_n)$ está normalmente distribuido con promedio nulo o bien que X tiene una densidad normal, si su densidad es una función de la forma

$$f(x) = c \exp \left[-\frac{1}{2} x^t A x \right]$$

donde c representa una constante y A una matriz simétrica. Si esto ocurre y si B es una matriz no singular, entonces el vector $Y = XB$ tiene la densidad

$$c |\det B|^{-1} \exp \left[-\frac{1}{2} y^t B^{-1} A B^{-1} y \right]$$

que es nuevamente una densidad normal. Además, es bien sabido que podemos elegir como B una matriz ortogonal, tal que $B^{-1} A B^{-1} = {}^t B A B = D$ sea una matriz diagonal:

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{pmatrix}$$

con esta elección de B, el vector $Y = (Y_1, \dots, Y_n)$ tiene la densidad

$$c \prod_{k=1}^n \exp \left[-\frac{1}{2} \lambda_k y_k^2 \right]$$

Para que esta función sea integrable sobre R^n , los números λ_k deben ser todos positivos. Por consiguiente, las variables que componen el vector Y son mutuamente independientes e Y_k tiene la densidad

$$N(0, \lambda_k^{-1}) = \frac{\sqrt{\lambda_k}}{2\pi} \exp \left[-\frac{1}{2} \lambda_k y_k^2 \right],$$

de donde resulta

$$\text{Cov } Y = D^{-1} = {}^t_B A^{-1} B$$

y también

$$c = \frac{(\lambda_1 \dots \lambda_n)^{1/2}}{(2\pi)^{n/2}} = \frac{(\det D)^{1/2}}{(2\pi)^{n/2}} = \frac{(\det A)^{1/2}}{(2\pi)^{n/2}}$$

Por otro lado, puesto que $X = Y {}^t_B$,

$$\text{Cov } X = B(\text{Cov } Y) {}^t_B = A^{-1},$$

de donde se deduce que escribiendo

$$W = \text{Cov } X,$$

la densidad del "vector normal" X es

$$(3.30) \text{ en } f(x) = \frac{(\det W)^{-1/2}}{(2\pi)^{n/2}} \exp \left[-\frac{1}{2} x {}^t W^{-1} x \right].$$

Más generalmente, si $a \in R^n$, la densidad $f(x-a)$, donde f es la función (3.30), se llama la densidad normal con promedio a y matriz de covarianzas W y se denota por $N(a; W)$.

De (3.30) se deduce que si el vector aleatorio $X = (X_1, \dots, X_n)$ está normalmente distribuido, la condición necesaria y suficiente para que las variables X_i sean mutuamente independientes es que $W = \text{Cov } X$ sea una matriz diagonal.

Un vector aleatorio $X = (X_1, \dots, X_n)$ formado por variables aleatorias independientes con la misma densidad φ sobre R se llama en Estadística una muestra de la densidad φ . En dichas condiciones, las variables aleatorias

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{y} \quad S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

se llaman, respectivamente, la media muestral y la varianza muestral.

Ahora estamos en condiciones de enunciar el resultado más importante de esta sección, a saber:

(3.31) TEOREMA (R. Fischer). Si X es una muestra de una densidad normal $N(\mu, \sigma^2)$, entonces la media y la varianzas muestrales son independientes.

DEMOSTRACION. Las variables aleatorias

$$Y_i = \frac{X_i - \mu}{\sigma} \quad (i = 1, 2, \dots, n)$$

son mutuamente independientes y tiene todas la misma densidad normal $N(0,1)$, lo que significa que el vector $Y = (Y_1, \dots, Y_n)$ es una muestra de $N(0,1)$ y tiene una densidad normal en R^n con promedio cero y matriz de covarianzas $\text{Cov } Y = I =$ la matriz unitaria.

De las relaciones $X_i = \sigma Y_i + \mu$ se deduce

$$\bar{X} = \sigma \bar{Y} + \mu,$$

donde $\bar{Y} = (1/n) \sum Y_i \sim N(0, 1/n)$ y por consiguiente,

$$(3.32) \quad \sqrt{n} \frac{\bar{X} - \mu}{\sigma} = \sqrt{n} \bar{Y} \sim N(0, 1).$$

Además, puesto que $X_i - \bar{X} = \sigma(Y_i - \bar{Y})$,

$$S^2 = (1/n) \sum (X_i - \bar{X})^2 = (\sigma^2/n) \sum (Y_i - \bar{Y})^2, \text{ de donde}$$

$$(3.33) \quad \frac{nS^2}{\sigma^2} = \sum_{i=1}^n (Y_i - \bar{Y})^2.$$

Consideremos ahora un cambio de variables $Z = YB$,

donde

$$B = \begin{pmatrix} 1/\sqrt{n} & * & \dots & * \\ 1/\sqrt{n} & * & \dots & * \\ \vdots & & \dots & \\ 1/\sqrt{n} & * & \dots & * \end{pmatrix} \quad (3.34)$$

es una matriz ortogonal con el único requisito de que todos los elementos de la primera columna valen $1/\sqrt{n}$; de modo que

$$Z_1 = (1/\sqrt{n}) \sum Y_i = \sqrt{n} \bar{Y}.$$

Puesto que B es ortogonal,

$$\begin{aligned} Z_1^2 + \dots + Z_n^2 &= \sum Y_i^2 = \sum [(Y_i - \bar{Y}) + \bar{Y}]^2 = \\ &= \sum [(Y_i - \bar{Y})^2 + 2\bar{Y}(Y_i - \bar{Y}) + \bar{Y}^2] = \\ &= \sum (Y_i - \bar{Y})^2 + n\bar{Y}^2 = \sum (Y_i - \bar{Y})^2 + Z_1^2. \end{aligned}$$

Luego,

$$\sum (Y_i - \bar{Y})^2 = Z_2^2 + \dots + Z_n^2;$$

es decir, en virtud de (3.32) y (3.33),

$$(3.34) \quad \sqrt{n} \frac{\bar{X} - \mu}{\sigma} = Z_1$$

y también

$$(3.35) \quad \frac{n S^2}{\sigma^2} = Z_2^2 + \dots + Z_n^2 .$$

Pero el vector Z tiene una densidad normal con promedio cero y matriz de covarianzas $\text{Cov } Z = {}^t B I B = I$, lo cual implica que sus componentes Z_i son mutuamente independientes, lo cual demuestra el teorema de Fischer. Más aún, la variable (3.35) tiene una densidad $\chi^2(n-1)$, de donde se deduce que la variable aleatoria

$$(3.36) \quad \sqrt{n-1} \frac{\bar{X} - \mu}{S}$$

tiene la densidad $t(n-1)$ que hemos calculado en la sección precedente.

Además de ser un teorema matemático interesante, el hecho de que (3.36) tenga una densidad que sólo depende del "tamaño" n de la muestra es fundamental en las aplicaciones estadísticas del Cálculo de Probabilidades.

THE UNIVERSITY OF CHICAGO

PHYSICS DEPARTMENT

CHICAGO, ILL.

OFFICE OF THE DEAN

Dear Sir:

I have the pleasure to inform you that your application for admission to the Ph.D. program in Physics has been reviewed and approved by the Department. You are invited to enroll in the fall semester of 1954. The Department is pleased to have you join our faculty and to contribute to our research program.

Sincerely,
[Signature]

Enclosed is a copy of the Department's letter to you regarding your admission to the Ph.D. program. Please contact the Department office if you have any questions.

I am sure that you will find the Department a most interesting and stimulating environment in which to pursue your studies. We look forward to your arrival in Chicago in the fall.

APENDICE A

§1. Construcción de ciertas funciones C^∞ .

Desde el punto de vista de la integración nada hay más sencillo que las funciones características de intervalos ó cubos en \mathbb{R}^n ; pero, como se trata de funciones discontinuas, resulta conveniente aproximarlas por funciones más regulares, por ejemplo continuas (fig.1).

Pero una función como la indicada en el gráfico no es suficientemente "suave" -requerimiento esencial en muchas cuestiones debido a los ángulos; "redondeando" éstos puede conseguirse una buena aproximación diferenciable e incluso - según veremos enseguida - infinitamente diferenciable.

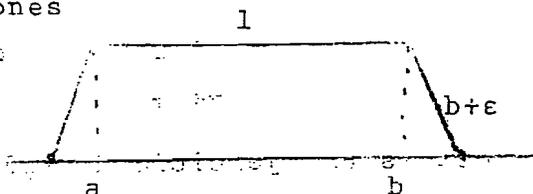


fig.1

A tal efecto, empezamos construyendo aplicaciones diferenciables (de clase C^∞) que se "pegan" bien en puntos tales como $a-\epsilon$ ó $b+\epsilon$ de la figura 1.

Lema A.1.1.

La función $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ definida por

$$\varphi(t) = \begin{cases} e^{-1/t} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

es no negativa, positiva para $t > 0$ y de clase C^∞ .

Demostración: Lo único no evidente es que φ es infinitamente diferenciable en $t=0$; como las derivadas laterales a izquierda de

todos los órdenes existen y son obviamente nulas, bastará probar que para todo $k \geq 0$ es

$$\varphi^{(k+1)}(+0) = \lim_{t \rightarrow 0} \frac{\varphi^{(k)}(t) - \varphi^{(k)}(0)}{t} = 0$$

La prueba es por inducción sobre $k \geq 0$ y se reduce a ver

que

$$\lim_{t \rightarrow 0} \frac{e^{-1/t}}{t^k} = 0$$

Poniendo $t = 1/x$, esto es consecuencia de

$$\lim_{x \rightarrow +\infty} x^k \cdot e^{-x} = 0$$

que es un ejercicio de aplicación de la regla de L'Hospital.

Veamos más consecuencias de A.1.1.: si $\alpha > 0$ definimos

$\psi_\alpha: \mathbb{R} \rightarrow \mathbb{R}$

$$\psi_\alpha(t) = \frac{\varphi(t/\alpha)}{\varphi(t/\alpha) + \varphi(1-t/\alpha)}$$

donde φ es la aplicación del lema anterior. Entonces es fácil

ver que

Lema A.1.2.

La aplicación ψ_α es no negativa, creciente, de clase C^∞ , nula para $t \leq 0$ e igual a 1 para $t \geq \alpha$ (fig. 2).

Obtenemos entonces (cuando $\alpha=1$):

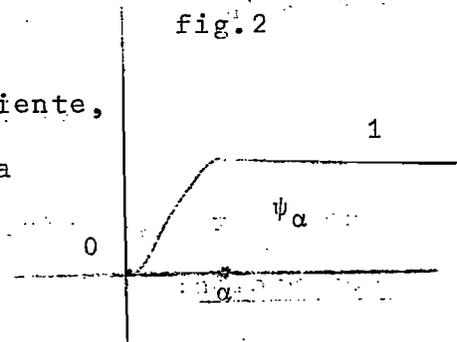


fig.2

Proposición A.1.3.

Si $a \in \mathbb{R}^n$, para todo $r > 0$ la aplicación $h_r: \mathbb{R}^n \rightarrow \mathbb{R}$ dada por

$$h_r(x) = \psi_1\left(2 + \frac{|x-a|}{r}\right) \cdot \psi_1\left(2 - \frac{|x-a|}{r}\right)$$

es no negativa, de clase C^∞ nula para $|x-a| \geq 2r$, menor que 1 si $r < |x-a| < 2r$ e igual a 1 si $|x-a| \leq r$.

Demostración: Se trata de una verificación; nótese que como la composición de funciones de clase C^∞ es de clase C^∞ , h_r resulta seguramente de clase C^∞ en $\mathbb{R}^n - \{a\}$ (donde $x \rightarrow |x-a|$ es de clase C^∞). Pero en un entorno de a (exactamente la bola de radio r y centro a) h_r es constante, y entonces de clase C^∞ en todo \mathbb{R}^n . Enseguida obtenemos el interesante resultado:

Proposición A.1.4.

Sea $K \subset \mathbb{R}^n$ un compacto y sea U un entorno de K . Existe entonces una aplicación $h: \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^∞ que verifica:

- $0 < h(x) < 1$ para todo $x \in \mathbb{R}^n$
- $h(x) = 1$ si $x \in K$
- $\text{sop}(h) \subset U$

Demostración: Por ser K compacto hay un conjunto finito a_1, \dots, a_m

de puntos de K y $r_1 > 0, \dots, r_m > 0$ tales que

$$K \subset \bigcup_{j=1}^m B(a_j, r_j) \quad \text{y} \quad K(a_j, 2r_j) \subset U \quad (1 \leq j \leq m)$$

Para cada uno de ellos hay (por A.1.3.) una aplicación $h_j: \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^∞ tal que $h_j(K(a_j, r_j)) = 1$, $0 < h_j(x) < 1$ si $x \in \mathbb{R}^n$ y $h_j(x) = 0$ si $x \notin B(a_j, 2r_j)$.

Si $f = \sum_{j=1}^m h_j$, f es de clase C^∞ no negativa, $f(x) > 0$ si $x \in K$ y $\text{sop}(f) \subset U$ $\text{sop}(h_j) \subset U$; sea $\alpha = \{\min_{x \in K} f(x)\}$. (Ciertamente $\alpha > 0$ y entonces $h = \psi_\alpha f$ (donde ψ_α está definida por (1)) verifica todas las condiciones a)b)c) requeridas (cf. A.1.2.)

Corolario A.1.5.

Sea $H \subset \mathbb{R}^n$ una hipersuperficie (o más generalmente una subvariedad diferenciable de \mathbb{R}^n), sea $K \subset H$ un compacto y V un entorno

de K en H . Existe entonces una aplicación $\psi: H \rightarrow \mathbb{R}$ de clase C^∞ que verifica:

a) $0 \leq \psi(x) \leq 1$ si $x \in H$

b) $\psi(x) = 1$ si $x \in K$

c) $\text{sop}(\psi) \subset V$

Demostración: Hay un entorno U de K en \mathbb{R}^n tal que $K \cup H \cap V$; entonces si $h: \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^∞ es dada por A.1.3., la aplicación $\psi = h/H$ responde a los requerimientos a)b)c).

§2. Algunos teoremas de topología

Supongamos que X es un espacio métrico (o más general: un espacio topológico, aunque en los desarrollos que siguen será suficiente considerar el caso en que X es un subconjunto de \mathbb{R}^n).

Una familia de subconjuntos $(A_i)_{i \in I}$ de X se dice localmente finita si para cada $x \in X$ hay un entorno U de x que corta solo a un número finito de conjuntos A_i , vale decir: para cada $x \in X$ hay un entorno U de x y un conjunto $J \subset I$ de índices

que verifican $U \cap A_i = \emptyset$ si $i \notin J$.

Muchas de las propiedades de las familias finitas pueden extenderse al caso de familias localmente finitas; por ejemplo:

Lema A.2.1.

Si $(A_i)_{i \in I}$ es una familia localmente finita, $\bigcup_{i \in I} A_i = \bigcup_{i \in I} \bar{A}_i$.

Demostración: Siempre es

$$\bigcup_{i \in I} \bar{A}_i \supseteq \bigcup_{i \in I} A_i$$

(pues el conjunto de la derecha es cerrado y contiene a todos los

A_i , por lo tanto contiene a todos los \bar{A}_i). Supongamos ahora que

x es adherente a $\bigcup_{i \in I} A_i$; por la definición, tendremos un entor-

no U de x que no corta a ningún A_i con $i \notin J$ ($J \subset I$ finito), así

que

$$U \cap \bigcup_{i \notin J} A_i = \emptyset \quad U \cap \bigcup_{i \in J} A_i \neq \emptyset.$$

Escribiendo $x \in \bigcup_{i \in J} A_i \cup \bigcup_{i \notin J} A_i$, vemos que necesariamente es

$x \in \bigcup_{i \in J} A_i$. Como esta unión es finita resulta entonces $x \in \bigcup_{i \in J} \bar{A}_i$,

así que seguramente es $x \in \bigcup_{i \in I} \bar{A}_i$.

Corolario A.2.2.

La unión de una familia localmente finita de conjuntos cerrados es un conjunto cerrado.

El siguiente resultado es válido en un contexto más general ("espacios topológicos normales") con demostración más engorrosa que la que damos aquí para un espacio métrico X :

Proposición A.2.3:

Sea $(U_i)_{i \in I}$ un cubrimiento abierto localmente finito de X . Hay entonces otro cubrimiento abierto $\bar{V}_i \in U_i$ para todo $i \in I$ (mismo conjunto de índices!) que verifica $\bar{V}_i \subset U_i$ para todo $i \in I$.

Demostración: Para cada $i \in I$ consideramos la función continua $\varphi_i: X \rightarrow \mathbb{R}$ definida por $\varphi_i(x) = d(x, X - U_i)$; claramente es $\varphi_i(x) \geq 0$ para todo $x \in X$, $\varphi_i(x) > 0$ para todo $x \in U_i$ (y $\varphi_i(x) = 0$ si $x \notin U_i$); como la familia es un cubrimiento, para cada $x \in X$ habrá por lo menos un índice $i \in I$ para el cual $\varphi_i(x) > 0$.

Pongamos $\varphi(x) = \sup \{ \varphi_i(x) : i \in I \}$; para cada x es $\varphi(x) > 0$. Afirmamos que φ es una función continua (y finita). En efecto, dado $x_0 \in X$ hay un entorno U de x_0 que corta solo a finitos U_i , y por lo tanto hay un $J \subset I$ finito tal que $\varphi_i|_U = 0$ si $i \notin J$.

Es evidente entonces que $(\varphi|_U)(x) = \sup \{ (\varphi_i|_U)(x) : i \in J \}$ que es una función continua finita. En consecuencia $\varphi|_U$ es continua, es decir: para cada $x_0 \in X$ hay un entorno U de x_0 tal que

$\varphi|_U: U \rightarrow \mathbb{R}$ es continua, y de aquí sigue sin dificultad la continuidad de φ .

Definimos ahora $V_i = \left\{ x : \varphi_i(x) > \frac{1}{2} \varphi(x) \right\}$ para cada $i \in I$; cada V_i resulta abierto (por ser φ continua). Además $\bar{V}_i \subset \{ x : \varphi_i(x) \geq \frac{1}{2} \varphi(x) \} \subset U_i$ para cada $i \in I$; para ver que efectivamente los V_i ($i \in I$) cubren X notemos que para un $x \in X$ hay sólo finitos i_1, \dots, i_r en I que verifican $\varphi_{i_1}(x) > 0, \dots, \varphi_{i_r}(x) > 0$. Si

los ordenamos por $0 < \psi_{i_1}(x) < \dots < \psi_{i_r}(x)$ es claro que $\psi_{i_r}(x) > \frac{1}{2} \psi_{i_1}(x) = \frac{1}{2} \psi(x)$, luego $x \in V_{i_r}$.

A los efectos de comprender cabalmente el enunciado (y la prueba) del importante resultado A.2.5. correspondi6 hacer m6s disquisici6nes sobre conjuntos compactos en R^n ; por de pronto si X es un subconjunto de R^n y $A \subset X$ es un subconjunto compacto de X (o sea: tiene la propiedad de Heine-Borel respecto de cubrimientos por abiertos en X) entonces A es un subconjunto compacto de R^n , y rec6procamente.

Pero, si K es un compacto en R^n , $K \cap X$ no necesariamente es compacto en X pues puede no ser compacto en R^n ; en particular $K(a,r) \cap X$ (con $r > 0$ y $a \in X$) puede no ser compacto. Ejemplo: $n=1$, $[0,1] \cap Q$ no es compacto ya que no es cerrado.

Obviamente esto depende de X , por lo que procedemos a fijar una nomenclatura para los subconjuntos que gozan de propiedades razonables en tal sentido: asi decimos que X es localmente cerrado si para cada $x \in X$ hay un $r_0 > 0$ tal que $K(x_0, r_0) \cap X$ es compacto.

Observaciones: A.2.4.

- a) Como $K(x, r_0) \cap X$ es acotado, la definici6n puede reformularse diciendo que para cada $x \in X$ hay un $\delta > 0$ tal que el conjunto $K(x, r_0) \cap X$ es cerrado.
- b) Con la notaci6n anterior, si $r > r_0$ resulta que $K(x, r) \cap X$ es cerrado (o compacto) pues es la intersecci6n del cerrado $K(x, r_0) \cap X$ con $K(x, r)$.

c) Todo subconjunto cerrado $X \subset \mathbb{R}^n$ es localmente cerrado (evidente).

d) Todo abierto $\Omega \subset \mathbb{R}^n$ es localmente cerrado; pues si $x \in \Omega$ hay un $\delta > 0$ tal que $B(x, \delta) \subset \Omega$. Luego si $0 < r_0 < \delta$ es claro que $K(x, r_0) \cap \Omega = K(x, r_0)$ bien cerrado.

e) Si X e Y son subconjuntos localmente cerrados también lo es su intersección; pues si $x \in X \cap Y$ y $K(x, r_0) \cap X$, $K(x, r_1) \cap Y$ son cerrados, también lo es su intersección, $K(x, r_2) \cap X \cap Y$ ($r_2 = \min(r_0, r_1)$).

f) Toda hipersuperficie- o más generalmente, toda subvariedad diferenciable en \mathbb{R}^n , es localmente cerrado; en efecto, dado $x \in X$ por definición hay un entorno abierto U de x en \mathbb{R}^n , un entorno abierto V de $0 \in \mathbb{R}^n$ y un difeomorfismo $h: U \rightarrow V$ tal que

$$h(U \cap X) = \{y \in V : y_{k+1} = 0, \dots, y_n = 0\}$$

(para hipersuperficie, $k = n-1$).

Entonces si $K(x, r_0) \subset U$ y $K = h(K(x, r_0)) \subset V$, K es compacto y $h(K(x, r_0) \cap X) = \{y \in K : y_{k+1} = 0, \dots, y_n = 0\}$ que es compacto, luego también $K(x, r_0) \cap X$ es compacto (ya que $h: K(x, r_0) \rightarrow K$ es homeomorfismo).

Finalmente veamos cuando un cubrimiento es más fino que otro:

si A y B son cubrimientos -de un conjunto cualquiera X - diremos que A es más fino que B , o que A es un refinamiento de B si cada $A \in A$ está contenido en algún $B \in B$ (y escribiremos $A > B$).

Ciertamente $A \subset B$ (o sea: A es un subcubrimiento de B) es condición suficiente para que $A > B$; pero ello no es de ningún modo necesario (tómese por ejemplo en $X = \mathbb{R}^n$, B = familia de todas las bolas abiertas, A = familia de todos los cubos cerrados. Claramente $A > B$ y $B > A$ pero $A \cap B = \emptyset$).

El siguiente resultado fundamental puede probarse bajo hipótesis más generales; pero el enunciado que damos es ampliamente suficiente para cubrir las aplicaciones usuales.

Teorema A.2.5.

Sea X un subconjunto de \mathbb{R}^n y sea B un cubrimiento abierto de X . Existe entonces un cubrimiento abierto A de X con las siguientes propiedades:

- a) A es numerable
- b) A es localmente finito
- c) A es más fino que B .

Si además X es localmente cerrado, A se puede determinar de manera que también verifique:

- d) Para cada $A \in A$, \bar{A} es compacto.

Demostración: Escribamos $B = (B_\lambda)_{\lambda \in L}$; por definición cada B_λ es de la forma $X \cap W_\lambda$, con cada W abierto en \mathbb{R}^n .

Para cada $x \in X$ y cada $\lambda \in L$ con $x \in B_\lambda$ hay un r (dependiente de x y λ) tal que $K(x, 2r) \subset W_\lambda$; si X es además localmente cerrado este r puede elegirse de manera que $K(x, 2r) \cap X$ sea compacto (cf. A.2.4.b)). La familia $B(x, r)$ correspondiente ciertamente cubre X , y entonces por la propiedad de Lindeloff hay una subfamilia numerable N de bolas abiertas $B(x_i, r_i)$ ($i \geq 1$) que también cubre X .

Evidentemente los conjuntos $B(x_i, 2r_i)$ ($i \geq 1$) también cubren X .

Definimos

$$A_1 = B(x_1, 2r_1) \cap X$$

$$A_{m+1} = [B(x_{m+1}, 2r_{m+1}) - \bigcup_{i=1}^m K(x_i, r_i)] \cap X$$

Los conjuntos A_k ($k \geq 1$) forman la familia A prometida; en primer lugar son evidentemente abiertos en X (son intersecciones de abiertos de \mathbb{R}^n con X) y la familia A es numerable por construcción.

Que A cubre X se ve así: supongamos que $x \in X$ y sea m el menor entero tal que $x \in B(x_m, 2r_m) \cap X$. Si $m=1$ resulta $x \in A_1$; si $m > 1$, no puede ser $x \in \bigcup_{i=1}^{m-1} B(x_i, 2r_i)$ y con mayor razón ser $x \notin \bigcup_{i=1}^{m-1} K(x_i, r_i)$. Luego (por definición) obtenemos $x \in A_m$.

probar que A es localmente finita, supongamos $x \in X$; ciertamente será $x \in U = B(x_k, r_k) \cap X$ para algún $k \geq 1$. Como $B(x_k, r_k) \cap A_m = \emptyset$ si $m > k$ vemos que U sólo puede cortar a A_1, \dots, A_m y entonces queda probado que A es localmente finita.

Finalmente siendo para todo $m \geq 1$

$$A_m \subset B(x_m, 2r_m) \cap X \subset K(x_m, 2r_m) \cap X$$

resulta que si X es localmente cerrado, y habiendo elegido los r de manera que $K(x, 2r) \cap X$ sea compacto resultará $A_m \subset K(x_m, 2r_m) \subset X$ así que A_m también será compacto.

El teorema queda completamente probado, ya que evidentemente cada A_m está contenido en $B(x_m, 25r_m) \cap X_m$ que por construcción está contenido en algún $W_\lambda \cap X = B_r$ (es decir, $A > B$).

§3. Particiones de la unidad.

Supongamos que X es un espacio métrico (o topológico) y que $\psi_i: X \rightarrow \mathbb{R}$ ($i \in I$) es una familia de aplicaciones tales que "la familia $\text{sop}(\psi_i)_{i \in I}$ es localmente finita" (1)

Entonces no hay ningún inconveniente en definir la aplicación

$$\sum_{i \in I} \psi_i: X \rightarrow \mathbb{R}$$

mediante la regla - para cada $x \in X$:

$$\left(\sum_{i \in I} \psi_i\right)(x) := \sum_{i \in I} \psi_i(x) \quad (2)$$

En efecto para cada $x \in X$ hay sólo un número finito de índices $i \in I$ para los cuales es $\psi_i(x) \neq 0$ y por lo tanto el miembro de la derecha en (2) es -para cada $x \in X$ - una suma finita (más precisamente: solo un número finito de sumandos son no nulos).

Es fácil ver que (siempre que se verifique la condición

(1)), estas sumas $\sum_{i \in I} \psi_i$ tienen las propiedades de las sumas

finitas; por ejemplo si $I = I_1 \cup I_2$ con $I_1 \cap I_2 = \emptyset$ vale

$$\sum_{i \in I_1} \psi_i + \sum_{i \in I_2} \psi_i = \sum_{i \in I} \psi_i$$

Lema A.3.1.

Sea X un espacio métrico y ψ_i ($i \in I$) una familia de aplicaciones de X en \mathbb{R} que verifica la condición (1). Entonces:

- a) Si todas las ψ_i son continuas también lo es $\sum_{i \in I} \psi_i$.
- b) Si X es una hipersuperficie en \mathbb{R}^n (o más generalmente, una variedad diferenciable) y todas las ψ_i son de clase C^k ($1 \leq k < \infty$), también lo es $\sum_{i \in I} \psi_i$.

Demostración: Si $a \in X$ hay un abierto U que contiene a a y una parte finita $J \subset I$ tales que $\text{sop}(\psi_i) \cap U = \emptyset$ si $i \notin J$; por con-

$$\left(\sum_{i \in I} \psi_i \right) | U = \sum_{i \in J} \psi_i | U$$

Como una suma finita de funciones continuas (resp: de clase C^k) es continua (resp: de clase C^k) sigue de la fórmula precedente que $\sum_{i \in I} \psi_i | U$ es continua. Como la continuidad (resp: de clase C^k), resultan a) y b). Veamos ahora un importante resultado que generaliza de manera notable la proposición A.1.4.:

Teorema A.3.2.

Sea $A \subset \mathbb{R}^n$ un conjunto cerrado y sea U un entorno abierto de A .

Existe entonces una aplicación $h: \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^∞ que verifica:

- a) $0 \leq h(x) \leq 1$ para todo $x \in \mathbb{R}^n$
- b) $h(x) = 1$ si $x \in A$
- c) $\text{sop}(h) \subset U$

Demostración: Apliquemos el teorema A.2.5. al cubrimiento

$B = \{U, X-A\}$ de $X = \mathbb{R}^n$; sea $A = \{U_j: j \geq 1\}$ verificando a)b)c)d) de dicho teorema. Como A es localmente finito podemos conseguir otro cubrimiento abierto V_j ($j \geq 1$) de tal manera que $\bar{V}_j \subset U_j$ para todo $j \geq 1$ (prop. A.2.3.).

Siendo cada \bar{V}_j compacto, de A.1.4. deducimos la existencia de aplicaciones $h_j: \mathbb{R}^n \rightarrow \mathbb{R}$ de clase C^∞ , $0 < h_j(x) < 1$ para cada $x \in \mathbb{R}^n$ de tal modo que $h_j(\bar{V}_j) = 1$ y $\text{sop}(h_j) \subset U_j$ (así que la familia de los soportes de las h_j es localmente finita).

Sea $J = \{i \geq 1: U_i \cap A \neq \emptyset\} = \{i \geq 1: U_i \subset U\}$ y pongamos $\varphi = \sum_{i \in J} h_i$, $\psi = \sum_{i \notin J} h_i$; ambas son aplicaciones de clase C^∞ (A.3.1.b) y como $\mathbb{R}^n = \bigcup_{i \geq 1} V_i$ resulta que $\varphi(x) + \psi(x) > 0$ para todo x . Si $h = \varphi / (\varphi + \psi)$ ciertamente h es de clase C^∞ y (verificada), también verifica b) puesto que si $x \in A$ necesariamente es $\psi(x) = 0$ (si $i \notin J$ es $U_i \cap A = \emptyset$, y en particular $\text{sop}(h_i) \cap A = \emptyset$).

Finalmente es $\text{sop}(h) = \text{sop}(\varphi) = \bigcup_{i \in J} \text{sop}(h_i) \subset \bigcup_{i \in J} U_i \subset U$ gracias a A.2.1.

Observación A.3.3. Si sólo se pretende que h sea sólo continua la prueba es bastante trivial: basta considerar un entorno abierto V de A con $\bar{V} \subset U$ (por ejemplo $V = \{x \in \mathbb{R}^n: d(x, A) < 1/2d(x, \mathbb{R}^n - U)\}$) y en seguida poner

$$h(x) = \frac{d(x, \mathbb{R}^n - V)}{d(x, A) + d(x, \mathbb{R}^n - V)}$$

notando que $\text{sop}(h) = \bar{V}$.

Ahora estamos en condiciones de demostrar teoremas de "partición de la unidad"; de manera general, si X es un espacio métrico (o topológico) llamamos partición de la unidad a toda familia $\varphi_i (i \in I)$ de aplicaciones continuas $\varphi_i: X \rightarrow \mathbb{R}$ que verifican

- i) $0 < \varphi_i(x)$ para todo $x \in X$
- ii) $\text{sop}(\varphi_i)$ ($i \in I$) es una familia localmente finita
- iii) $\sum_{i \in I} \varphi_i = 1$

Si X es hipersuperficie (o una variedad diferenciables) y las φ_i son de clase C^k , hablamos de partición de la unidad de clase C^k ($1 \leq k \leq \infty$) (cf. lema 2.16).

Teorema A.3.4. (Teorema de partición de la unidad")

Sea X una hipersuperficie (o una variedad diferenciable) y sea $B = (B_\lambda)_{\lambda \in L}$ un cubrimiento abierto de X .

- a) Existe una partición de la unidad numerable φ_i ($i \geq 1$) de clase C^∞ tal que
 - i) $\text{sop}(\varphi_i)$ es compacto para todo $i \geq 1$
 - ii) La familia $\text{sop}(\varphi_i)$ ($i \geq 1$) es un refinamiento de B .
- b) Existe una partición de la unidad h_λ ($\lambda \in L$) (mismo conjunto de índices!) tal que
 - i) $\text{sop}(h_\lambda) \subset B_\lambda$ para todo $\lambda \in L$
 - ii) $h_\lambda = 0$ salvo un conjunto numerable de índices $\lambda \in L$

Demostración: Para cada $\lambda \in L$ y cada $x \in B_\lambda$ hay una carta local (U, φ) alrededor de x tal que $\varphi(U) = \bigcup S \subset B_\lambda$ el cubrimiento abierto formado por tales S_i (indicando S) evidentemente es más fino que B . Aplicando A.2.5. obtenemos un cubrimiento abierto numerable $A = \{A_i : i \geq 1\}$ que refina a S (y por lo tanto a B), con cada \bar{A}_i compacto; y siendo A localmente finito podemos producir otro cubrimiento abierto V_i ($i \geq 1$) de modo que $\bar{V}_i \subset A_i$ (A.2.3.)

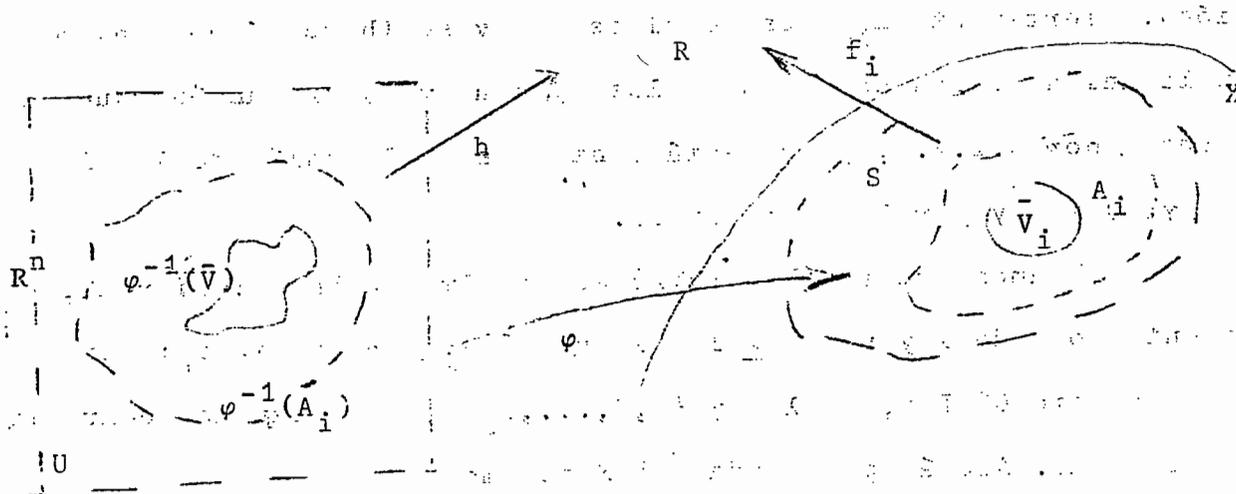
Ahora afirmamos que es posible construir (para cada $i \geq 1$) aplicación $f_i: X \rightarrow \mathbb{R}$ de clase C^∞ que verifique $f_i(X) \subset [0,1]$, $f_i(\bar{V}_i) = 1$ y $\text{supp}(f_i) \subset A_i$.

En efecto, hay un $S \in \mathcal{S}$ para el cual $A_i \subset S$, con carta local correspondiente $\varphi: U \rightarrow S$; tomando $\varphi^{-1}(A_i)$ abierto en \mathbb{R}^k , $\varphi^{-1}(\bar{V}_i)$ cerrado en \mathbb{R}^k y utilizando A.3.2. construimos una aplicación $h: \mathbb{R}^k \rightarrow \mathbb{R}$ de clase C^∞ , $h(\mathbb{R}^k) \subset [0,1]$, $h(\varphi^{-1}(\bar{V}_i)) = 1$ y $\text{supp}(h) \subset \varphi^{-1}(A_i) \subset U$.

Se pone entonces

$$f_i(x) = \begin{cases} 0 & \text{si } x \notin S \\ h \circ \varphi^{-1}(x) & \text{si } x \in S \end{cases}$$

y es inmediato que f_i tiene las propiedades requeridas (ver figura).



Como los \bar{V}_i ($i \geq 1$) cubren X (y forman una familia localmente finita) la función $f = \sum_{i \geq 1} f_i$ es de clase C^∞ (A.3.1.b)) y $f(x) > 0$ para todo $x \in X$, así que para cada $i \geq 1$ la función $g_i = f_i/f$

es de clase C^∞ , $g_i \geq 0$ y $\text{sop}(g_i) = \text{sop}(f_i) \subset A_i$.

Como cada \bar{A}_i es compacto, es $\text{sop}(g_i)$ compacto para todo $i \geq 1$; es obvio que $\sum_{i \geq 1} g_i = 1$ y que la familia $\text{sop}(g_i)$ ($i \geq 1$) refina a S y por ende a B .

La parte a) del teorema queda completamente demostrada.

Para la parte b), elegimos para cada $i \geq 1$ un $\lambda \in L$ de manera que $A_i \subset B_{\lambda_i}$, $A_2 \subset B_{\lambda_2}$, ... En otros términos: sea $\gamma: \mathbb{N} \rightarrow L$ una aplicación que verifica $A_i \subset B_{\gamma(i)}$ para cada $i \geq 1$ (una tal aplicación γ existe por ser A_i ($i \geq 1$) más fino que B).

Se pone:

$$h_\lambda = \sum_{i \in C_\lambda} g_i$$

donde $C_\lambda = \{i \geq 1 : \gamma(i) = \lambda\} = \gamma^{-1}(\lambda)$ (Si $C_\lambda = \emptyset$, $h_\lambda = 0$ por definición). Ciertamente $h_\lambda \geq 0$ es de clase C^∞ y $\text{sop}(h_\lambda)$ está contenido en la unión de los soportes de las g_i con $\gamma(i) = \lambda$, unión que es cerrada por A.2.2., y que está contenida en la unión de los A_i con $\gamma(i) = \lambda$ y por lo tanto en B_λ .

Veamos que $\text{sop}(h_\lambda)$ ($\lambda \in L$) es localmente finita: si $x \in X$, hay un entorno U de x y un $i_0 \geq 1$ tal que $U \cap A_i = \emptyset$ si $i > i_0$; por consiguiente $U \cap B_\lambda = \emptyset$ si $\lambda \notin \gamma\{1, 2, \dots, i_0\}$, lo que queda demostrado. Además para cada $x \in X$ tenemos

$$1 = \sum_{i \geq 1} g_i(x) = \sum_{\lambda \in L} \left(\sum_{i \in C_\lambda} g_i(x) \right) = \sum_{\lambda \in L} h_\lambda(x)$$

Finalmente, $h_\lambda = 0$ si $\lambda \notin \gamma(\mathbb{N})$; esto concluye la demostración del teorema.

Observación A.3.5. Si sólo pretende una partición de la unidad con funciones continuas se puede hacer una prueba más sencilla en particular en lo que se refiere a la construcción de las f_i de la parte a). Pues con un razonamiento idéntico al utilizado en A.3.3. se obtiene f_i que vale 1 sobre \bar{V}_i , con soporte en A_i . El resto de la demostración no cambia prácticamente (excepto que la continuidad de f apela a A.3.1.a)).

Ejercicio A.3.5. Demostrar A.3.2. para el caso "A cerrado es una hipersuperficie (o variedad diferenciable) X " (Sug: tomar una partición de la unidad como en el caso A.3.4.b)) para

$$S = \{X-A, U\}.$$

Ejercicio A.3.7. Sea $A \subset \mathbb{R}^m$ y sea $\varphi: A \rightarrow \mathbb{R}^n$ de clase C^k (ver definición en el capítulo II, §1). Demostrar que hay un abierto Ω que es entorno de A y una $F: \Omega \rightarrow \mathbb{R}^n$ de clase C^k tal que $F|_A = \varphi$ (sug: con notación de §1, capítulo II tomar Ω como la unión de todos los U y $F = \sum_U g_U \cdot f_U$ para una partición de la unidad g_U ($U \in \mathcal{U}$)).

Referencias

- [1] BLUMENSON, L.E., A derivation of n-dimensional spherical coordinates. Amer. Math. Monthly, Vol.67,1960, páginas 63-66.
- [2] FAVA, N.A., On a formula related to polar coordinates in R^n . Mathematicae Notae, Año XXVI, 1977/78, páginas 47-49.
- [3] FEDERER, H., Geometric measure theory. Springer-Verlag, New York, 1969, página 250.
- [4] SCHWARTZ, J., The formula for changing variables in a multiple integral. Amer.Math. Monthly 61 (1954), páginas 81-85.
- [5] SANTALO, L.A. , Introduction to integral geometry. Actualités Scientifiques et Industrielles 1198, Hermann, Paris, 1952.
- [6] SANTALO, L.A., Integral geometry and geometric probability. Addison-Wesley, Reading, 1976.
- [7] SPIVAK, M., Calculus on manifolds. Benjamin , New York, 1965.
- [8] WHEEDEN, R.L. y ZYGMUND, A., Measure and integral. Marcel Dekker, New York, 1977.
- [9] ZAAANEN, A., An introduction to the theory of integration. North Holland, Amsterdam, 1961.