Capítulo 5

Funciones diferenciables

Funciones de varias variables: Derivada según un vector y derivadas parciales. Aplicaciones diferenciables. Condición suficiente de diferenciabilidad. Reglas del cálculo diferencial. Regla de la cadena. Espacio tangente. Gradiente.

En este capítulo comienza el cálculo diferencial para funciones de varias variables reales con valores escalares o vectoriales. Aunque la noción elemental de derivada parcial es insuficiente para poder desarrollar con ella un cálculo diferencial satisfactorio, similar al de funciones de una variable, sin embargo la sola consideración de las derivadas parciales tiene aplicaciones interesantes que se exponen en la primera sección: Las funciones con derivadas parciales acotadas son continuas y las funciones con derivadas parciales nulas son constantes si su dominio es un abierto conexo.

Una de las primeras aplicaciones de la noción de derivada parcial la proporciona la condición necesaria de extremo relativo. Con este modesto recurso ya se pueden abordar y resolver algunos problemas de optimización.

Sin embargo la sola consideración de las derivadas parciales no proporciona un marco satisfactorio para desarrollar un cálculo diferencial con resultados similares al caso de las funciones de una variable. El marco adecuado lo proporciona la noción de aplicación diferenciable. En este caso la diferencial en un punto proporciona una aproximación local de la función mediante un polinomio de primer grado (lo que equivale a una aproximación local, mediante una aplicación lineal, del incremento de la función). Esta noción, para funciones de dos variables, se motiva con el problema de definir planos tangentes a 'superficies' dadas en forma explícita $z = f(x, y)$.

Después de introducir la noción de aplicación diferenciable y de haber estudiado las condiciones necesarias para la diferenciabilidad se demuestra la condición suficiente de diferenciabilidad en términos de continuidad de las derivadas parciales. Este resultado proporciona una herramienta eficaz para obtener la diferenciabilidad de funciones en las situaciones habituales.

A continuación se establecen las reglas usuales del cálculo diferencial (diferencial de la suma, del producto, etc) haciendo especial énfasis en la regla de la cadena, en la que se apoyan la mayor parte de los resultados útiles del cálculo diferencial.

Después de introducir la matriz Jacobiana y las notaciones habituales del cálculo diferencial, se consideran las funciones con valores reales y para ellas, aprovechando la estructura euclídea de $\mathbb{R}^n$, se expresa la diferencial en términos del vector gradiente. La noción de gradiente, que se formula de modo intrínseco, permite nuevas interpretaciones geométricas y físicas, y con las que se aprecia su interés desde el punto de vista de las aplicaciones.

Finaliza el capítulo con la existencia de espacio tangente a la gráfica de una aplicación diferenciable en un punto. También se considera el problema de la existencia de espacio tangente a conjuntos de nivel de aplicaciones diferenciables (en particular, para superficies dadas en forma implícita) y para conjuntos que se pueden parametrizar mediante una aplicación diferenciable (en particular, para superficies dadas en forma paramétrica). Para ello merece la pena anticipar algunos resultados posteriores sobre existencia de funciones implícitas, con el fin de que al finalizar el capítulo quede clara la noción precisa de espacio tangente y se conozcan los métodos para calcularlo en las situaciones habituales.

En este capítulo se formulan las primeras definiciones en el contexto general de los espacios normados sobre el cuerpo $\mathbb{R}$, es decir, para funciones $\mathbf{f} : \Omega \to F$, definidas en un abierto Ω de un espacio normado $(E, \| \ \|)$ con valores en otro espacio normado $(F, \| \ \|)$. (Para simplificar la notación designamos igual las normas de E y F). Este contexto, además de proporcionar mayor generalidad, obliga a usar una notación más compacta que, incluso en el caso $E = \mathbb{R}^n$, $F = \mathbb{R}^m$, evita el uso engorroso de coordenadas en aquellas cuestiones en las que estas no desempeñan un papel especial. Como los principales resultados, aplicaciones e interpretaciones geométricas, se refieren casi siempre al caso $E = \mathbb{R}^n$, $F = \mathbb{R}^m$, el lector que lo desee puede considerar, desde ahora en adelante, esta situación particular.

5.1. Derivada según un vector

Es natural que el estudio local, en el entorno de un punto, de una función de varias variables se apoye en la consideración de las funciones de una variable que se obtienen restringiendo la función a las rectas que pasan por el punto. Esta idea es la que inspira la siguiente definición.

Definición 5.1 *Sea* $\mathbf{f} : \Omega \to F$ *una aplicación definida en un abierto* Ω *del espacio normado* $(E, \| \ \|)$*, con valores en el espacio normado* $(F, \| \ \|)$*. Dados* $\mathbf{a} \in \Omega$ *y* $\mathbf{v} \in E$*, la derivada de* $\mathbf{f}$ *en* $\mathbf{a}$ *según el vector* $\mathbf{v}$ *es el vector de* F *dado por*

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{t \to 0} \frac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t}$$

en el supuesto de que el límite exista.

La derivada según un vector unitario se llama derivada direccional: Si $\|\mathbf{v}\| = 1$*, la derivada direccional de* $\mathbf{f}$ *en* $\mathbf{a}$ *según la dirección* $\mathbf{v}$ *es* $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$*.*

Cuando $E = \mathbb{R}^n$, las derivadas según los vectores de la base canónica

$$\mathbf{e}_1 = (1, 0, \cdots 0), \ \mathbf{e}_2 = (0, 1, 0, \cdots 0), \ \mathbf{e}_n = (0, 0, \cdots, 1)$$

tienen un interés especial: La derivada $D_{\mathbf{e}_1}\mathbf{f}(\mathbf{a})$, si existe, viene dada por

$$\lim_{t \to 0} = \frac{\mathbf{f}(\mathbf{a}+t\mathbf{e}_1)-\mathbf{f}(\mathbf{a})}{t} = \lim_{t \to 0} \frac{\mathbf{f}(a_1+t, a_2, \cdots, a_n) - \mathbf{f}(a_1, a_2, \cdots, a_n)}{t}$$

Es decir, $D_{\mathbf{e}_1}\mathbf{f}(\mathbf{a})$ es la derivada en a_1 de la función parcial $x_1 \to \mathbf{f}(x_1, a_2, a_3, \cdots a_n)$, definida en un entorno abierto de a_1. Análogamente, $D_{\mathbf{e}_j}\mathbf{f}(\mathbf{a})$ es la derivada, en a_j, de la función parcial $x_j \to \mathbf{f}(a_1, \cdots a_{j-1}, x_j, a_{j+1}, \cdots a_n)$.

Por esta razón a la derivada $D_{\mathbf{e}_j}\mathbf{f}(\mathbf{a})$, cuando existe, se le llama derivada parcial de $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x_1, x_1, \cdots x_n)$, respecto a la variable x_j, en el punto $\mathbf{a}$, para la que se usan las notaciones habituales

$$D_j\mathbf{f}(\mathbf{a}); \qquad \frac{\partial \mathbf{f}}{\partial x_j}(\mathbf{a}).$$

Para funciones de dos o tres variables, es habitual llamar x, y, z a las variables x_1, x_2, x_3 y suele resultar cómodo emplear las notaciones

$$\mathbf{f}_x(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial x}(\mathbf{a}); \quad \mathbf{f}_y(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial y}(\mathbf{a}); \quad \mathbf{f}_z(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial z}(\mathbf{a}).$$

La ventaja de las derivadas parciales respecto a las derivadas según vectores arbitrarios reside en que su cálculo suele ser una tarea mecánica basada en las reglas usuales del cálculo de derivadas de funciones reales de una variable real.

Veremos en la siguiente sección que para las funciones diferenciables siempre existe la derivada según un vector $\mathbf{v} = (v_1, v_2, \cdots v_n)$, que se puede calcular en términos de las derivadas parciales usando la fórmula

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = v_1 D_1 f(\mathbf{a}) + v_2 D_2 \mathbf{f}(\mathbf{a}) + \cdots + v_n D_n \mathbf{f}(\mathbf{a})$$

La siguiente proposición nos dice que para funciones con valores en $\mathbb{R}^m$ el cálculo de la derivada según un vector, se puede realizar componente a componente:

Proposición 5.2 *Si* $F = \mathbb{R}^m$ *y* $\mathbf{f} = (f_1, f_2, \cdots, f_m)$, *la derivada* $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ *existe si y sólo si existen las derivadas de las componentes,* $D_{\mathbf{v}}f_1(\mathbf{a})$, $D_{\mathbf{v}}f_2(\mathbf{a})$, ... , $D_{\mathbf{v}}f_m(\mathbf{a})$ *y en ese caso* $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = (D_{\mathbf{v}}f_1(\mathbf{a}), D_{\mathbf{v}}f_2(\mathbf{a}), \cdots, D_{\mathbf{v}}f_m(\mathbf{a}))$.

Dem: Basta aplicar la proposición 4.3 a la función $\boldsymbol{\gamma}_{\mathbf{v}}(t) = \mathbf{f}(\mathbf{a}+t\mathbf{v})$ cuyas componentes son las funciones $f_j(\mathbf{a}+t\mathbf{v})$, $1 \le j \le m$. ■

Interpretación física de la derivada según un vector:
Cuando $E = \mathbb{R}^3$ y $F = \mathbb{R}$, es posible dar diversas interpretaciones físicas de la derivada según un vector: Supongamos que en un recinto Ω del espacio la temperatura T (o cualquier otra magnitud física, como la presión) depende del punto $(x, y, z) \in \Omega$ y viene dada mediante la función $T = f(x, y, z)$. Si una partícula recorre una línea recta que pasa por $\mathbf{p} \in \Omega$ con velocidad constante $\mathbf{v}$ entonces $D_{\mathbf{v}}f(\mathbf{p})$ representa la razón de cambio de la temperatura de la partícula respecto al tiempo, cuando la partícula pasa por $\mathbf{p}$: En efecto, si suponemos que la partícula pasa por $\mathbf{p}$ en el instante $t = 0$, un poco después, en el instante $t > 0$ la partícula

se encuentra en la posición $\mathbf{p} + t\mathbf{v}$, luego $[f(\mathbf{p} + t\mathbf{v}) - f(\mathbf{p})]/t$ es la razón de cambio de la temperatura durante el intervalo de tiempo $(0, t)$. Pasando al límite cuando $t \to 0$ se obtiene que $D_{\mathbf{v}} f(\mathbf{p})$ es la razón de cambio instantánea en $t = 0$, es decir, cuando la partícula pasa por $\mathbf{p}$.

Interpretación geométrica de la derivada según un vector:
La interpretación geométrica de la derivada según un vector se basa en la siguiente

Definición 5.3 *Sea M un subconjunto no vacío del espacio normado F y $\mathbf{p} \in M$. Si existe una aplicación $\boldsymbol{\gamma} : (-\epsilon, \epsilon) \to F$, derivable en $t = 0$, tal que*

$$\boldsymbol{\gamma}(t) \in M \text{ si } |t| < \epsilon, \quad \boldsymbol{\gamma}(0) = \mathbf{p}, \quad y \quad \boldsymbol{\gamma}'(0) = \mathbf{v}$$

diremos que $\mathbf{v} \in F$ es un vector tangente a M en $\mathbf{p}$. En lo que sigue $T_{\mathbf{p}}(M)$ designará el conjunto de los vectores tangentes a M en $\mathbf{p}$.

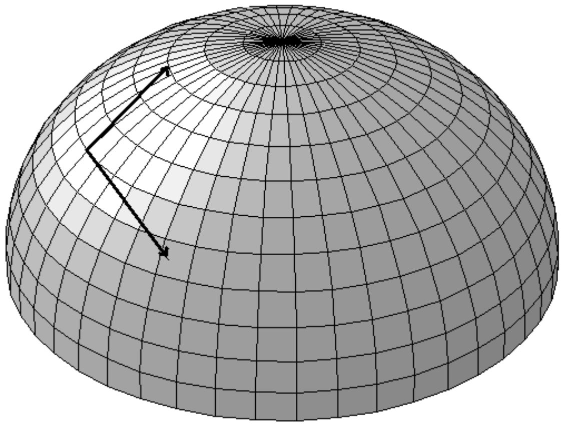
En las condiciones de la definición 5.1, $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \boldsymbol{\gamma}'_{\mathbf{v}}(0)$ donde $\boldsymbol{\gamma}_{\mathbf{v}}(t) = \mathbf{f}(\mathbf{a} + t\mathbf{v})$ está definida en un entorno de 0, luego $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ es un vector tangente a $\mathbf{f}(\Omega)$ en $\mathbf{p} = \mathbf{f}(\mathbf{a})$. Si existe $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ es fácil ver que también existe $D_{\mathbf{u}} f(\mathbf{a})$ donde $\mathbf{u} = t\mathbf{v}$, con $t \in \mathbb{R}$, y que su valor es $D_{\mathbf{u}}\mathbf{f}(\mathbf{a}) = tD_{\mathbf{v}}\mathbf{f}(\mathbf{a})$. Es decir, cuando $\mathbf{u} = t\mathbf{v}$ recorre una recta que pasa por 0, los vectores tangentes $D_{\mathbf{u}}\mathbf{f}(\mathbf{a})$ forman un subespacio vectorial unidimensional. Si estos vectores se dibujan con origen en $\mathbf{p}$, sus extremos recorren una recta afín, y diremos que esta recta es tangente a $\mathbf{f}(\Omega)$ en $\mathbf{p}$.

a) Comencemos considerando el caso de una 'superficie' paramétrica, $M = \boldsymbol{\varphi}(\Omega)$ descrita como imagen de una aplicación $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^3$, definida en un abierto $\Omega \subset \mathbb{R}^2$. Las coordenadas (x, y, z) de un punto genérico de la 'superficie' son variables que dependen de dos parámetros independientes $(s, t) \in \Omega$. Si $\varphi_1, \varphi_2, \varphi_3$ son las componentes de $\boldsymbol{\varphi}$, se suele decir que $M = \boldsymbol{\varphi}(\Omega)$ es una 'superficie' de ecuaciones paramétricas $x = \varphi_1(s, t)$, $y = \varphi_2(s, t)$, $z = \varphi_3(s, t)$.

Dado un punto $\mathbf{a} \in \Omega$, uno de los objetivos del cálculo diferencial es el de definir de manera razonable, cuando sea posible, el plano tangente a la 'superficie' paramétrica $M = \boldsymbol{\varphi}(\Omega)$ en un punto $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ de la misma. La idea básica es que este plano debe estar formado por rectas que pasan por $\mathbf{p}$ y son tangentes a 'curvas' contenidas en la superficie. Dado un vector $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$, una 'curva' de este tipo es la que tiene la ecuación paramétrica $\boldsymbol{\gamma}_{\mathbf{v}}(t) = \boldsymbol{\varphi}(\mathbf{a} + t\mathbf{v})$, donde t es un parámetro real que varía en el entorno de 0, $A_{\mathbf{v}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{v} \in \Omega\}$.

La derivada $\boldsymbol{\gamma}'_{\mathbf{v}}(0) = D_{\mathbf{v}}\boldsymbol{\varphi}(\mathbf{a})$, si existe, es un vector tangente a la 'superficie' paramétrica $M = \boldsymbol{\varphi}(\Omega)$ en el punto $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$, que se suele representar gráficamente mediante una flecha con origen en $\mathbf{p}$ y extremo en $\mathbf{p} + D_{\mathbf{v}}\boldsymbol{\varphi}(\mathbf{a})$. En particular, las derivadas parciales $D_1\boldsymbol{\varphi}(\mathbf{a})$, $D_2\boldsymbol{\varphi}(\mathbf{a})$, si existen, son vectores tangentes a M en $\mathbf{p}$.

Más adelante estudiaremos las condiciones que garantizan que $T_{\mathbf{p}}(M)$ es un espacio vectorial de dimensión 2, con base $\{D_1\boldsymbol{\varphi}(\mathbf{a}), D_2\boldsymbol{\varphi}(\mathbf{a})\}$.



Para ver un ejemplo muy concreto, consideremos la parametrización habitual de la esfera unidad, $\mathbf{r}(\varphi,\theta) = (\cos\varphi\cos\theta, \cos\varphi\operatorname{sen}\theta, \operatorname{sen}\varphi)$, donde $\varphi \in (-\pi/2, \pi/2)$, $\theta \in (0, 2\pi)$, son respectivamente, la latitud y la longitud geográficas del punto $\mathbf{r}(\varphi,\theta)$. Los vectores

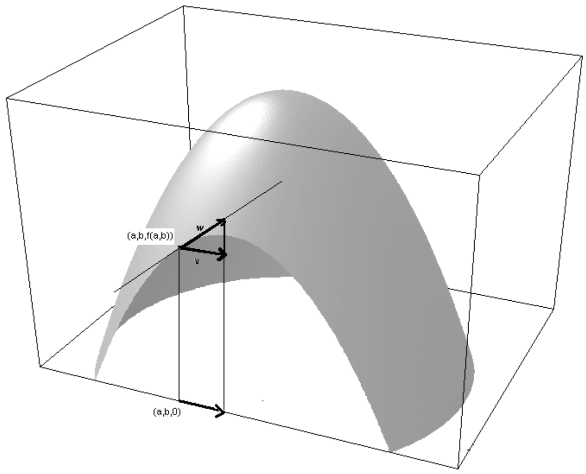
$$\frac{\partial \mathbf{r}}{\partial \varphi}(\varphi,\theta) = (-\operatorname{sen}\varphi\cos\theta, -\operatorname{sen}\varphi\operatorname{sen}\theta, \cos\varphi)$$

$$\frac{\partial \mathbf{r}}{\partial \theta}(\varphi,\theta) = (-\cos\varphi\operatorname{sen}\theta, \cos\varphi\cos\theta, 0)$$

son tangentes, respectivamente, al meridiano y al paralelo de la esfera que pasa por el punto $\mathbf{r}(\varphi,\theta)$. Pronto podremos justificar, en este caso particular, que el conjunto de vectores tangentes a la esfera en el punto $\mathbf{r}(\varphi,\theta)$ forma un espacio vectorial de dimensión 2, generado por los vectores $\{D_1\mathbf{r}(\varphi,\theta), D_2\mathbf{r}(\varphi,\theta)\}$.

b) Consideremos ahora el caso de una función real de dos variables reales $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^2$ (caso $E = \mathbb{R}^2$, $F = \mathbb{R}$). Su gráfica $G(f) = \{(x, y, f(x, y)) : (x, y) \in \Omega\}$ es un subconjunto de $\mathbb{R}^3$ que admite una representación tridimensional como una 'superficie', sobre el plano xy. Se trata de una 'superficie' de un tipo muy especial, pues cada recta paralela al eje z, trazada por un punto $(x, y, 0)$, con $(x, y) \in \Omega$ la corta solamente en el punto $(x, y, f(x, y))$. Para un punto genérico $\mathbf{p} = (x, y, z) \in G(f)$, las dos primeras coordenadas, x, y, son variables independientes que recorren Ω, mientras que la tercera z, depende de ellas a través de la función $z = f(x, y)$. Estas 'superficies' son un caso particular de las consideradas en a) pues admiten la parametrización $\boldsymbol{\varphi}(x, y) = (x, y, f(x, y))$.

No obstante merece la pena ver, en este caso, la interpretación geométrica de la derivada direccional: Si $\mathbf{u} \in \mathbb{R}^2$ es unitario, $\|\mathbf{u}\| = 1$, ahora la derivada $D_{\mathbf{u}}f(a, b)$ es un número que representa la pendiente de la tangente en el punto $\mathbf{p} = (a, b, f(a, b))$ a la curva que se obtiene como intersección de la gráfica $G(f)$ con el plano que pasa por $(a, b, f(a, b))$ y $(a, b, 0)$ y es paralelo al vector $\mathbf{v} = (u_1, u_2, 0)$. (Obsérvese que esta recta tangente tiene la dirección del vector $\mathbf{w} = (u_1, u_2, D_{\mathbf{u}}f(a, b))$ pues en el triángulo rectángulo que tiene como cateto horizontal el vector unitario $\mathbf{v}$, y la hipotenusa en la recta tangente, la longitud del cateto vertical es la pendiente de la tangente, es decir $D_{\mathbf{u}}f(a, b)$).



En particular el número $D_1 f(a,b)$ (resp. $D_2 f(a,b)$) da la pendiente de la tangente a la curva que se obtiene cortando la gráfica de f con el plano que pasa por $\mathbf{p}$ y es paralelo al plano xz (resp. al plano yz).

Si se parametriza la 'superficie' $G(f)$ en la forma estandar, $G(f) = \boldsymbol{\varphi}(\Omega)$, con $\boldsymbol{\varphi}(x,y) = (x,y,f(x,y))$, se tiene $\varphi_1(x,y) = x$, $\varphi_2(x,y) = y$, $\varphi_3(x,y) = f(x,y)$, luego $D_{\mathbf{u}}\varphi_1(a,b) = u_1$, $D_{\mathbf{u}}\varphi_2(a,b) = u_2$, y según 5.2 volvemos a obtener que $D_{\mathbf{u}}\boldsymbol{\varphi}(a,b) = (u_1, u_2, D_{\mathbf{u}} f(a,b))$ es un vector tangente a $G(f)$ en el punto $\mathbf{p} = (a,b,f(a,b))$.

El siguiente ejemplo muestra una función $f : \mathbb{R}^2 \to \mathbb{R}$ que no es continua en $(0,0)$ pero existen las derivadas $D_{\mathbf{v}} f(0,0)$ según todos los vectores $\mathbf{v} \in \mathbb{R}^2$. Así se pone de manifiesto que la noción elemental de derivada según un vector es insuficiente para desarrollar con ella un cálculo diferencial satisfactorio similar al ya conocido para funciones reales de una sola variable.

Ejemplo 5.4 *La función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x,y) = \frac{xy^2}{x^2 + y^4} \quad \textit{si } (x,y) \neq (0,0), \quad f(0,0) = 0$$

no es continua en $(0,0)$, y para todo $\mathbf{v} \in \mathbb{R}^2$ existe $D_{\mathbf{v}} f(0,0)$. Además, las derivadas parciales $D_1 f(x,y)$, $D_2 f(x,y)$ no están acotadas en ningún entorno de $(0,0)$.

Dem: En $(0,0)$ existe la derivada según cualquier vector $\mathbf{v} = (v_1, v_2)$ pues en el caso no trivial $(v_1, v_2) \neq (0,0)$, es evidente que

$$\frac{f((0,0) + t(v_1, v_2)) - f(0,0)}{t} = \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4}$$

tiene límite cuando $t \to 0$ (si $v_1 = 0$ el límite es 0, y si $v_1 \neq 0$ el límite es v_2^2/v_1). Por otra parte, como $f(0,0) = 0$ y en todo entorno de $(0,0)$ hay puntos (de la parábola $x = y^2$) donde el valor de f es $1/2$ se sigue que f no es continua en $(0,0)$.

Para $y \in \mathbb{R}$, fijo la función de variable real $x \to f(x, y)$ es derivable en todo $x \in \mathbb{R}$, y la derivada vale

$$D_1 f(x, y) = \frac{y^6 - x^2 y^2}{(x^2 + y^4)^2}$$

Para $x \in \mathbb{R}$, fijo la función de variable real $y \to f(x, y)$ es derivable en todo $y \in \mathbb{R}$, y la derivada vale

$$D_2 f(x, y) = \frac{2x^3 y - 2xy^5}{(x^2 + y^4)^2}$$

En el punto $(0, 0)$, usando la definición es claro que existen las derivadas parciales

$$D_1 f(0, 0) = \lim_{h \to 0} [f(h, 0) - f(0, 0)]/h = 0; D_2 f(0, 0) = \lim_{k \to 0} [f(0, k) - f(0, 0)]/k = 0$$

$D_1 f(x, y)$ no está acotada en ningún entorno de $(0, 0)$ porque $D_1 f(0, y) = 1/y^2$. $D_2 f(x, y)$ no está acotada en ningún entorno de $(0, 0)$ porque en los puntos de parábola $x = my^2$, con $(m \neq 1)$, toma los valores $D_1 f(my^2, y) = c(m)/y$, donde $c(m) = 2(m^3 - m)/(m^2 + 1)^2 \neq 0$. ■

Aunque la sola consideración de las derivadas parciales es insuficiente para desarrollar un cálculo diferencial satisfactorio hay algunas aplicaciones interesantes de esta noción que recogemos a continuación.

Algunas aplicaciones de la noción de derivada parcial. La función considerada en el ejercicio 5.4 también sirve para mostrar que la acotación de las derivadas parciales es una hipótesis esencial para la validez del siguiente resultado, que proporciona una condición suficiente para la continuidad.

Proposición 5.5 *Sea $\Omega \subset \mathbb{R}^n$ abierto y $\mathbf{f} : \Omega \to F$ una aplicación tal que en cada punto $\mathbf{x} \in \Omega$ existen las derivadas parciales $D_1\mathbf{f}(\mathbf{x})$, $D_2\mathbf{f}(\mathbf{x})$, $\cdots$, $D_n\mathbf{f}(\mathbf{x})$. Si las derivadas parciales $D_k\mathbf{f}(\mathbf{x})$, $1 \leq k \leq n$ están acotadas entonces $\mathbf{f}$ es continua. Si además $\Omega = \mathbb{R}^n$ entonces $\mathbf{f}$ es uniformemente continua.*

DEM: a) Por hipótesis existe $M > 0$ tal que $\|D_k\mathbf{f}(\mathbf{x})\| \leq M$ para todo $\mathbf{x} \in \Omega$ y todo $k \in \{1, 2, \cdots, n\}$. Dado un punto $\mathbf{a} \in \Omega$, sea $r > 0$ tal que $B(\mathbf{a}, r) \subset \Omega$. Para cada $\mathbf{x} \in B(\mathbf{a}, r)$ el segmento de extremos $(x_1, x_2, \cdots, x_n)$, $(a_1, x_2, \cdots, x_n)$ está contenido en $B(\mathbf{a}, r)$ y aplicando el corolario 4.8 a la función de una variable real $\boldsymbol{\varphi}_1(t) = \mathbf{f}(t, x_2, x_3, \cdots, x_n)$ en el segmento $[a_1, x_1]$ se obtiene

(1) $\|\mathbf{f}(x_1, x_2, \cdots, x_n) - \mathbf{f}(a_1, x_2, \cdots, x_n)\| \leq M|x_1 - a_1|$

El segmento de extremos $(a_1, x_2, \cdots, x_n)$, $(a_1, a_2, x_3, \cdots, x_n)$ también está contenido en $B(\mathbf{a}, r)$, y aplicando el teorema del incremento finito a la función de una variable real $\boldsymbol{\varphi}_2(t) = \mathbf{f}(a_1, t, x_3, \cdots, x_n)$ en el segmento $[a_2, x_2]$ se obtiene

(2) $\|\mathbf{f}(a_1, x_2, \cdots, x_n) - \mathbf{f}(a_1, a_2, \cdots, x_n)\| \leq M|x_2 - a_2|$

Análogamente se acota cada sumando de la suma telescópica

$$\begin{aligned}\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) = \quad & \mathbf{f}(x_1, x_2, x_3, \cdots, x_n) - \mathbf{f}(a_1, x_2, x_3, \cdots, x_n) \quad + \\ & \mathbf{f}(a_1, x_2, x_3, \cdots, x_n) - \mathbf{f}(a_1, a_2, x_3 \cdots, x_n) \quad + \\ & \cdots\cdots\cdots \quad + \\ & \mathbf{f}(a_1, a_2 \cdots a_{n-1}, x_n) - \mathbf{f}(a_1, a_2, a_3, \cdots, a_n)\end{aligned}$$

y utilizando la desigualdad triangular se obtiene

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq M \sum_{j=1}^{n} |x_j - a_j| = M \|\mathbf{x} - \mathbf{a}\|_1$$

de donde se sigue que $\mathbf{f}$ es continua en $\mathbf{a}$.
b) Cuando $\Omega = \mathbb{R}^n$, con el mismo razonamiento se obtiene que para todo par de puntos $\mathbf{a}, \mathbf{x} \in \mathbb{R}^n$ se verifica

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq M \|\mathbf{x} - \mathbf{a}\|_1$$

y de aquí se sigue que f es uniformemente continua. ■

NOTA: El resultado de la proposición 5.5 se puede localizar en un punto $\mathbf{a}$: Si sólo se supone que en una bola $B(\mathbf{a}, r) \subset \Omega$ existen todas las derivadas parciales y están acotadas, entonces $\mathbf{f}$ es continua en $B(\mathbf{a}, r)$, y en particular es continua en $\mathbf{a}$.

Por otra parte, la última afirmación de la proposición 5.5 no es cierta cuando $\Omega \neq \mathbb{R}^n$: En el ejercicio 5.35 veremos un ejemplo de una función $g : \Omega \to \mathbb{R}$, que no es uniformemente continua en $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$, pero tiene derivadas parciales acotadas en Ω.

Para el siguiente resultado se usa la noción de conjunto conexo que se estudia en los cursos de Topología General.

Proposición 5.6 *Sea $\Omega \subset \mathbb{R}^n$ un abierto conexo y $\mathbf{f} : \Omega \to F$ tal que en cada $\mathbf{x} \in \Omega$ existen y son nulas las derivadas parciales $D_1\mathbf{f}(\mathbf{x})$, $D_2\mathbf{f}(\mathbf{x})$, $\cdots$, $D_n\mathbf{f}(\mathbf{x})$. Entonces $\mathbf{f}$ es constante.*

DEM: En virtud de la proposición 5.5 $\mathbf{f}$ es continua. Más aún, la demostración de esta proposición, con $M = 0$, conduce a que $\mathbf{f}$ es constante en cada bola $B(\mathbf{a}, r) \subset \Omega$.

El conjunto no vacío $A = \{\mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a})\}$ es cerrado para la topología relativa de Ω porque $\mathbf{f}$ es continua. Basta ver que A también es un subconjunto abierto del abierto conexo Ω para concluir que $A = \Omega$, y con ello que $\mathbf{f}$ es constante. Efectivamente, si $\mathbf{b} \in A$ existe $\rho > 0$ tal que $B(\mathbf{b}, \rho) \subset \Omega$, y en esta bola $\mathbf{f}$ es constante, luego $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{b})$ para todo $\mathbf{x} \in B(\mathbf{b}, r)$. Como $\mathbf{b} \in A$ se cumple que $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$, luego $B(\mathbf{b}, \rho) \subset A$ y con esto queda establecido que A es abierto. ■

Como aplicación de la proposición 5.6 mostramos en el siguiente ejemplo una alternativa para definir en $\Omega = \mathbb{R}^2 \setminus \{(x, 0) : x \leq 0\}$ la función argumento principal que asigna a cada $(x, y) \in \Omega$ el único argumento $\theta = f(x, y)$ del número complejo $x + iy$ que cumple $-\pi < \theta < \pi$. La definición utiliza la determinación principal $\operatorname{Arc\,tg} : \mathbb{R} \to (-\pi/2, \pi/2)$ de la función multivaluada arctan.

Ejemplo 5.7 *En* $\Omega_1 = \{(x,y) : y > 0\}$, $\Omega_2 = \{(x,y) : x > 0\}$, $\Omega_3 = \{(x,y) : y < 0\}$ *se definen, respectivamente, las funciones*

$$f_1(x,y) = \text{Arc tg}(-x/y) + \pi/2$$

$$f_2(x,y) = \text{Arc tg}(y/x),$$

$$f_3(x,y) = \text{Arc tg}(-x/y) - \pi/2$$

Entonces en el abierto $\Omega = \mathbb{R}^2 \setminus \{(x,0) : x \leq 0\} = \Omega_1 \cup \Omega_2 \cup \Omega_3$, *hay definida una función* $f : \Omega \to \mathbb{R}$, *con derivadas parciales continuas, tal que* $f|_{\Omega_i} = f_i$.

DEM: Para $i = 1, 2, 3$ y todo $(x,y) \in \Omega_i$, con un cálculo rutinario se obtiene

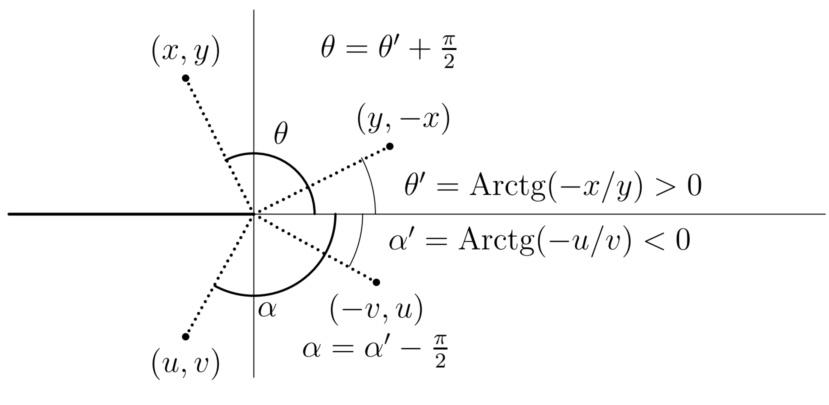
$$D_1 f_i(x,y) = \frac{-y}{x^2+y^2}; \quad D_2 f_i(x,y) = \frac{x}{x^2+y^2}$$

Según 5.6, $f_1 - f_2$ es constante en el abierto conexo $\Omega_1 \cap \Omega_2 = \{(x,y); x > 0, y > 0\}$ y su valor constante es $f_1(1,1) - f_2(1,1) = (-\pi/4 + \pi/2) - \pi/4 = 0$.

Análogamente, $f_2 - f_3$ es constante en $\Omega_2 \cap \Omega_3 = \{(x,y); x > 0, y < 0\}$ y su valor constante es $f_2(1,-1) - f_3(1,-1) = -\pi/4 - (\pi/4 - \pi/2) = 0$.

En definitiva, f_1 y f_2 coinciden en la intersección de sus dominios, y lo mismo ocurre con f_2 y f_3 (los dominios de f_1 y f_3 son disjuntos). Esto permite definir en $\Omega = \cup_{j=1}^{3} \Omega_j = \mathbb{R}^2 \setminus \{(x,0) : x \leq 0\}$ una función f cuya restricción a cada Ω_j sea f_j. Las derivadas parciales de esta función siguen siendo

$$D_1 f(x,y) = \frac{-y}{x^2+y^2}; \quad D_2 f(x,y) = \frac{x}{x^2+y^2}$$



Según la figura, se tiene

$$0 < \theta = f(x,y) = \text{Arc tg}(-x/y) + \pi/2 = \theta' + \pi/2 < \pi$$

$$-\pi < \alpha = f(u,v) = \operatorname{Arc\,tg}(-u/v) - \pi/2 = \alpha' - \pi/2 < 0$$

La función $f : \Omega \to \mathbb{R}$ tiene un significado geométrico muy claro: $\theta = f(x,y)$ es el único argumento del número complejo $(x+iy)$ que cumple $-\pi < \theta < \pi$. ■

Una de las aplicaciones más notables de la noción de derivada parcial la proporciona la condición necesaria de extremo relativo para funciones reales de varias variables reales:

Definición 5.8 *Una función $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, presenta en $\mathbf{a} \in \Omega$ un mínimo relativo (resp. mínimo relativo estricto) si existe una bola $B(\mathbf{a}, r) \subset \Omega$, tal que cuando $\|\mathbf{x} - \mathbf{a}\| < r$ (resp. $0 < \|\mathbf{x} - \mathbf{a}\| < r$) se verifica $f(\mathbf{a}) \leq f(\mathbf{x})$ (resp. $f(\mathbf{a}) < f(\mathbf{x})$).*

La definición de máximo relativo (resp. máximo relativo estricto) es análoga pero usando la desigualdad $f(\mathbf{a}) \geq f(\mathbf{x})$ (resp. $f(\mathbf{a}) > f(\mathbf{x})$).

Si f presenta en $\mathbf{a}$ un máximo o un mínimo relativo se dice que f presenta en $\mathbf{a}$ un extremo relativo.

Es obvio que f presenta en $\mathbf{a} \in \Omega$ un máximo relativo (resp. máximo relativo estricto) si y sólo si $-f$ presenta en $\mathbf{a} \in \Omega$ un mínimo relativo (resp. mínimo relativo estricto). A veces, en estas definiciones, en vez del calificativo "relativo" se emplea el de "local", y se habla de máximos locales, mínimos locales, extremos locales, etc.

Proposición 5.9 *Si $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, presenta en $\mathbf{a} \in \Omega$ un extremo local y existe la derivada $D_{\mathbf{u}}f(\mathbf{a})$ entonces $D_{\mathbf{u}}f(\mathbf{a}) = 0$. En particular, si existen las derivadas parciales $D_1f(\mathbf{a})$, $D_2f(\mathbf{a})$,.. $D_nf(\mathbf{a})$, todas son nulas.*

Dem: La función real de variable real $\varphi(t) = f(\mathbf{a} + t\mathbf{u})$ está definida en un entorno de $t = 0$ donde es derivable:

$$\varphi'(0) = \lim_{t \to 0} \frac{\varphi(t) - \varphi(0)}{t} = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t} = D_{\mathbf{u}}f(\mathbf{a})$$

Si f presenta en $\mathbf{a}$ un máximo (resp. mínimo) relativo entonces φ presenta en 0 un máximo (resp. mínimo) relativo y por lo tanto $D_{\mathbf{u}}f(\mathbf{a}) = \varphi'(0) = 0$. ■

Si $f : \Omega \to \mathbb{R}$ posee derivadas parciales en todo $\mathbf{x} \in \Omega$, se llaman *puntos estacionarios* de f, a las soluciones (si las hay) del sistema de n ecuaciones

$$D_1f(x_1, x_2, \cdots x_n) = 0, \quad D_2f(x_1, x_2, \cdots x_n) = 0, \ \cdots D_nf(x_1, x_2, \cdots x_n) = 0$$

Según la proposición 5.9 los puntos estacionarios de las funciones que tienen derivadas parciales en todos los puntos de su dominio son los puntos donde estas funciones pueden presentar extremos relativos. Con el siguiente ejemplo se muestra que puede haber puntos que sean estacionarios pero no sean puntos de extremo local

Ejemplo 5.10 *El origen $(0,0)$ es punto estacionario de $f(x,y) = (y - x^2)(y - 2x^2)$, pero f no presenta en $(0,0)$ un extremo relativo:*

DEM: Obsérvese que $f(0,0) = 0$ y que en todo entorno de $(0,0)$ hay puntos (x_1, y_1), (x_2, y_2) tales que $f(x_1, y_1) < f(0,0) < f(x_2, y_2)$. (Basta tomar $x_1^2 < y_1 < 2x_1^2$, $2x_2^2 < y_2$). ■

Aplicaciones al cálculo de extremos absolutos.
A veces ocurre que por algún razonamiento de tipo topológico tenemos la certeza de que una función continua $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, alcanza un máximo absoluto en algún punto $\mathbf{a} \in \Omega$. Si f, además de ser continua, tiene derivadas parciales en todo $\mathbf{x} \in \Omega$ entonces $\mathbf{a} \in E$, donde E es el conjunto de los puntos estacionarios de f. Si este conjunto es finito y se puede calcular explícitamente, quedará resuelto del problema de obtener el máximo absoluto de f, que vendrá dado por $\max\{f(\mathbf{e}) : \mathbf{e} \in E\}$.

Cuando alguna de las derivadas parciales de f no existe en algún punto, y también es finito el conjunto $H \subset \Omega$ formado por los puntos donde esto ocurre, para obtener el máximo absoluto basta calcular $\max\{f(\mathbf{e}) : \mathbf{e} \in E \cup H\}$. El mismo método se puede seguir para calcular el mínimo absoluto de una función continua, cuando se sabe de antemano que la función alcanza en algún punto de su dominio un mínimo absoluto. En los ejercicios resueltos 5.37, 5.38, se pueden ver ejemplos típicos del uso de este método.

Los ejercicios resueltos 5.39, 5.40 son ejemplos típico de problemas de extremos condicionados: Extremos de una función cuyas variables están sometidas a una condición (llamada condición de ligadura). Aunque el tratamiento general de este tipo de problemas se expone en el capítulo 9, aquí se muestra como algunos problemas de este tipo se pueden resolver con los recursos elementales de este capítulo.

Con los recursos disponibles en este momento también se puede abordar el problema de calcular los extremos absolutos de una función continua de dos variables sobre un compacto $K \subset \mathbb{R}^2$, siempre que la función tenga derivadas parciales en los puntos del interior de K y la frontera ∂K esté formada por una cantidad finita de curvas con una parametrización sencilla.

La estrategia para calcular los extremos absolutos es la siguiente: Un resultado general de topología asegura que f alcanza en K sus extremos absolutos. Un punto (a, b) donde $f|_K$ alcanza un extremo absoluto puede estar en el interior de K o en su frontera. En el primer caso (a, b) será un punto estacionario de f, es decir, será solución de las ecuaciones $D_1 f(x, y) = 0$, $D_2 f(x, y) = 0$. Si sabemos resolver estas ecuaciones y en el interior de K sólo hay un conjunto finito S de soluciones, cada punto de S será candidato a punto de extremo absoluto.
En el segundo caso, cuando $(a, b) \in \partial K$, es claro que $f|_{\partial K}$ alcanza un extremo absoluto en (a, b). Si la frontera ∂K se compone de varios arcos de curva con parametrizaciones sencillas, usando estas parametrizaciones, el cálculo de los extremos absolutos de $f|_{\partial K}$ se transforma en varios problemas de optimización unidimensional que se pueden abordar con las técnicas usuales del cálculo de funciones de una variable real. Los tres casos considerados en el ejercicio 5.41 se han resuelto siguiendo esta estrategia.

5.2. Aplicaciones diferenciables

Para motivar la noción de diferencial de una aplicación volvemos a considerar el problema de formular una definición razonable de plano tangente a una 'superficie' explícita $z = f(x,y)$ en un punto $\mathbf{p} = (a, b, f(a, b))$ de la misma.

Para lograr una definición razonable se debe cumplir que para cada $\mathbf{v} \in \mathbb{R}^2$ exista la derivada $D_{\mathbf{v}} f(a, b)$, pero ya hemos visto que este requisito es demasiado débil pues ni siquiera garantiza la continuidad de la función en el punto $\mathbf{a}$ (véase el ejercicio 5.4) Incluso cuando se supone que f continua, no se puede asegurar que el conjunto de vectores tangentes a $G(f)$ en $\mathbf{p} = (a, b, f(a, b))$ forme un plano, como ocurre en la función que interviene en el siguiente ejercicio

Ejemplo 5.11 *La función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x,y) = \frac{x^2 y}{x^2 + y^2} \quad \text{si } (x,y) \neq (0,0), \quad f(0,0) = 0$$

es continua en $(0,0)$ y para todo $\mathbf{v} \in \mathbb{R}^2$ existe $D_{\mathbf{v}} f(0,0) = f(\mathbf{v})$, y en el punto $\mathbf{p} = (0,0,0)$ de la gráfica $M = G(f)$ se cumple

$$T_{\mathbf{p}}(M) = \{(v_1, v_2, f(v_1, v_2)) : (v_1, v_2) \in \mathbb{R}^2\}$$

DEM: La continuidad de f en $(0,0)$ es consecuencia inmediata de la desigualdad $|f(x,y)| \leq |y|$ que se cumple en todo punto $(x,y) \in \mathbb{R}^2$.

En $(0,0)$ existe la derivada según cualquier vector $\mathbf{v} = (v_1, v_2)$ pues en el caso no trivial $(v_1, v_2) \neq (0,0)$, es evidente que

$$\frac{f((0,0) + t(v_1, v_2)) - f(0,0)}{t} = \frac{v_1^2 v_2}{v_1^2 + v_2^2}$$

luego existe $D_{\mathbf{v}} f(0,0) = f(\mathbf{v})$.

Para establecer la igualdad $T_{\mathbf{p}}(M)) = \{(v_1, v_2, f(v_1, v_2)) : (v_1, v_2) \in \mathbb{R}^2\}$ basta demostrar que todo $\mathbf{u} = (u_1, u_2, u_3) \in T_{\mathbf{p}}(M)$ satisface $u_3 = f(u_1, u_2)$, es decir $u_3(u_1^2 + u_2^2) = u_1^2 u_2$. Según la definición de vector tangente existe $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^3$, $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, derivable en $t = 0$ tal que $\gamma_3(t)(\gamma_1(t)^2 + \gamma_2(t)^2) = \gamma_1(t)^2 \gamma_2(t)$, $\gamma_j(0) = 0$, y $\gamma_j'(0) = u_j$ para $1 \leq j \leq 3$. En virtud de la definición de derivada

$$\gamma_j(t) = t[u_j + \epsilon_j(t)], \quad \text{donde} \quad \lim_{t \to 0} \epsilon_j(t) = 0$$

Se sigue que

$$(u_3 + \epsilon_3(t))[(u_1 + \epsilon_1(t))^2 + (u_2 + \epsilon_2(t))^2] = (u_1 + \epsilon_1(t))^2 (u_2 + \epsilon_2(t))$$

y pasando al límite cuando $t \to 0$ se obtiene que $u_3(u_1^2 + u_2^2) = u_1^2 u_2$. ■

Dejando aparte las consideraciones geométricas, para la función del ejercicio 5.11 la derivada según vectores $\mathbf{v} \to D_{\mathbf{v}} f(\mathbf{a}) = f(\mathbf{v})$ no proporciona nada nuevo. No se consigue así otra función más sencilla (lineal) que aproxime localmente a f.

Vemos así que para lograr una definición satisfactoria de plano tangente a la gráfica $G(f)$ en un punto $\mathbf{p} = (a, b, f(a, b)) \in G(f)$ debemos conseguir que $\{(\mathbf{v}, D_{\mathbf{v}} f(\mathbf{a})) : \mathbf{v} \in \mathbb{R}^2\}$ sea un subespacio vectorial de dimensión 2 y que se cumpla la igualdad

$$\{(\mathbf{v}, D_{\mathbf{v}} f(\mathbf{a})) : \mathbf{v} \in \mathbb{R}^2\} = T_{\mathbf{p}}(G(f))$$

Es claro que una condición necesaria y suficiente para ello es que la aplicación $\mathbf{v} \to D_{\mathbf{v}} f(a, b)$ sea lineal. En este caso, $T_{\mathbf{p}}(G(f)) \subset \mathbb{R}^3$ es un candidato a plano tangente. Sin embargo, para la definición satisfactoria de aplicación diferenciable en un punto $\mathbf{a}$ hay que exigir algo más que la continuidad de f en $\mathbf{a}$ y la existencia y linealidad de la aplicación $\mathbf{v} \to D_{\mathbf{v}} f(\mathbf{a})$ (véanse los ejemplos 5.15 y 5.42).

Definición 5.12 *Sea $\mathbf{f} : \Omega \to F$ definida en un abierto Ω de un espacio normado $(E, \|\ \|)$, con valores en otro espacio normado $(F, \|\ \|)$ y $\mathbf{a} \in \Omega$. Se dice que $\mathbf{f}$ es diferenciable en $\mathbf{a}$ cuando existe una aplicación lineal continua $L : E \to F$ que verifica*

$$\lim_{h \to 0} \frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ y se cambian las normas de E y F por otras equivalentes, es fácil ver que $\mathbf{f}$ sigue siendo diferenciable en $\mathbf{a}$, con la misma diferencial.

Obsérvese que hay una única aplicación lineal L que cumpla la condición de la definición anterior: Si $T \in \mathcal{L}(E, F)$ es otra aplicación lineal que cumple

$$\lim_{h \to 0} \frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - T(\mathbf{h})}{\|\mathbf{h}\|} = 0 \quad \text{resulta} \quad \lim_{h \to 0} \frac{T(\mathbf{h}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

y considerando el límite a través de $\{t\mathbf{v} : t > 0\}$ con $\mathbf{v} \in E$, $\mathbf{v} \neq 0$, se obtiene

$$0 = \lim_{t \to 0+} \frac{T(t\mathbf{v}) - L(t\mathbf{v})}{t\,\|\mathbf{v}\|} = \frac{T(\mathbf{v}) - L(\mathbf{v})}{\|\mathbf{v}\|} \quad \text{luego} \quad T(\mathbf{v}) = L(\mathbf{v}).$$

Definición 5.13 *En las condiciones de la definición 5.12, la única aplicación lineal continua $L : E \to F$ que interviene en ella se llama diferencial de $\mathbf{f}$ en el punto $\mathbf{a}$ y se denota $d\mathbf{f}(\mathbf{a})$. Si $\mathbf{f}$ es diferenciable en cada $\mathbf{a} \in \Omega$ se dice que $\mathbf{f}$ es diferenciable en Ω. En este caso $d\mathbf{f} : \Omega \to \mathcal{L}(E, F)$ es la aplicación que asigna a cada $\mathbf{a} \in \Omega$ la diferencial $d\mathbf{f}(\mathbf{a})$.*

La condición de que $\mathbf{f}$ sea diferenciable en $\mathbf{a}$ se expresa escribiendo

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\|\,\boldsymbol{\rho}(\mathbf{h}) \quad \text{con} \quad \lim_{\mathbf{h} \to 0} \boldsymbol{\rho}(\mathbf{h}) = 0$$

El significado de la noción de aplicación diferenciable es el siguiente: La aplicación $\mathbf{f}$ es diferenciable en $\mathbf{a}$ cuando es posible obtener una aproximación lineal $L(\mathbf{h})$ del incremento $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$ con la propiedad de que el error de la aproximación $\boldsymbol{\alpha}(\mathbf{h}) = [\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})] - L(\mathbf{h})$ sea pequeño frente a $\mathbf{h}$, en el sentido de que el error relativo $\boldsymbol{\alpha}(\mathbf{h}) / \|\mathbf{h}\|$ tiende hacia 0 cuando $\mathbf{h}$ tiende hacia 0. Si $\boldsymbol{\alpha}(\mathbf{h})$ es una función de $\mathbf{h} \in E$, definida en un entorno de 0, se suele escribir $\boldsymbol{\alpha}(\mathbf{h}) = o(\|\mathbf{h}\|)$ para indicar que se cumple

$$\lim_{\mathbf{h} \to 0} \frac{\boldsymbol{\alpha}(\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Con esta notación la condición de que $\mathbf{f}$ sea diferenciable en $\mathbf{a}$ se expresa en la forma más breve

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = o(\|\mathbf{h}\|)$$

Si dos funciones $\mathbf{f}$, $\mathbf{g}$, definidas en un entorno de $\mathbf{a}$, con $\mathbf{f}(\mathbf{a}) = \mathbf{g}(\mathbf{a})$, verifican la condición $\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{a}\|)$ se suele decir que $\mathbf{f}$ y $\mathbf{g}$ presentan en $\mathbf{a}$ una tangencia de primer orden, y también que $\mathbf{g}$ es una aproximación local de primer orden de $\mathbf{f}$ en el punto $\mathbf{a}$. Según esta terminología, cuando $\mathbf{f}$ es diferenciable en $\mathbf{a}$, la aplicación lineal $L(\mathbf{h})$ es una aproximación local de primer orden del incremento $\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})$ (en $\mathbf{h} = 0$) y la aplicación afín $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{x} - \mathbf{a})$ es una aproximación local de primer orden de $\mathbf{f}$ en el punto $\mathbf{a}$.

Comentarios sobre la definición de aplicación diferenciable. Con la noción de diferencial en un punto se persigue (igual que con la derivada, en el caso de funciones de una sola variable) una aproximación local de la función mediante un polinomio de primer grado. En el caso de las funciones de dos variables la aproximación local será de la forma $p(x, y) = Ax + By + C$. Cuando esto se formula en términos de los incrementos $h = x - a$, $k = y - b$, la diferencial proporciona una aplicación lineal $L(h, k) = Ah + Bk$ que, en un entorno de $(0, 0)$, aproxima localmente al incremento $f(a+h, b+k)-f(a, b)$. La noción de aproximación local que interviene en la definición de función diferenciable 5.12 exige que la aproximación local de $L(h, k) = Ah + Bk$ al incremento $f(a + h, b + k) - f(a, b)$ sea uniforme en todas las direcciones. Esta noción es la que permite desarrollar un cálculo diferencial satisfactorio, análogo al de funciones de una variable. Entre otras cosas permitirá demostrar la existencia de plano tangente a la gráfica de una función diferenciable de dos variables reales.

Para pensar en una situación concreta, que termine de aclarar el significado de la diferenciabilidad sigamos considerando el caso de una función de dos variables reales $z = f(x, y)$ que deseamos evaluar numéricamente en un punto (a, b) que sólo se puede manejar en forma aproximada (esto ocurrirá cuando a, b sean irracionales o procedan de medir magnitudes físicas). Supongamos además que sólo aceptamos resultados con error menor que $\epsilon = 10^{-12}$ (el tamaño que ya no distingue nuestra calculadora) y que f es diferenciable en (a, b). En este caso la aplicación lineal $L = df(a, b) : \mathbb{R}^2 \to \mathbb{R}$ es de la forma $L(h, k) = Ah + Bk$.

Según la definición de límite existe $r > 0$ tal que si $\max\{|h|, |k|\} < r$ se cumple $[f(a + h, b + k) - f(a, b) - L(h, k)] < \epsilon \max\{|h|, |k|\}$. Es decir, para incrementos que cumplen $\max\{|h|, |k|\} < r$, si utilizamos $f(a, b) + L(h, k)$ como una aproximación de $f(a + h, b + k)$ estaremos seguros de que cometemos un error menor que $10^{-12} \max\{|h|, |k|\}$, que es prácticamente despreciable frente a $\max\{|h|, |k|\}$.

Por otra parte, si utilizamos un ordenador para visualizar la gráfica de $z = f(x, y)$ y la de su aproximación local $z = f(a, b) + A(x - a) + B(y - b)$ y hacemos zoom en el punto $\mathbf{p} = (a, b, f(a, b))$ hasta que el entorno de $\mathbf{p}$ de radio r ocupe toda la pantalla, podremos ver que, dentro de este entorno, no se aprecia diferencia entre el trozo de plano y el trozo de la gráfica de f ya que el tamaño de la pantalla del ordenador es $\simeq r$, mientras que la diferencia de cotas entre un punto del plano y un punto de la gráfica, es menor que $10^{-12}r$.

Ejemplos

i) Si $\mathbf{f}$ es constante, es obvio que $\mathbf{f}$ es diferenciable en todo $\mathbf{a} \in \Omega$ y su diferencial es la aplicación lineal nula $\mathbf{0} \in \mathcal{L}(E, F)$.

ii) Si $L : E \to F$ es lineal continua entonces L es diferenciable en todo punto $\mathbf{x} \in E$ y su diferencial es constante: $dL(\mathbf{x}) = L$ para todo $\mathbf{x} \in E$.

En particular, si $E = \mathbb{R}^2$, $F = \mathbb{R}$ y $L(x, y) = Ax + By$, para todo $(a, b) \in \mathbb{R}^2$

$$dL(a, b)(h, k) = Ah + Bk$$

iii) Cuando $E = \mathbb{R}$, es fácil ver que $\mathbf{f}$ es derivable en a si y sólo si es diferenciable en a y en ese caso la diferencial es a la aplicación lineal $d\mathbf{f}(a) : \mathbb{R} \to F$, definida por $d\mathbf{f}(a)(h) = \mathbf{f}'(a)h$.

NOTA: En iii) $\mathbf{f}'(a)$ es un vector y h un número, y aparece escrito $\mathbf{f}'(a)h$ en lugar de $h\mathbf{f}'(a)$. Este convenio de notación para la ley externa de los espacios vectoriales es útil para lograr mayor analogía con las fórmulas habituales del cálculo diferencial de las funciones escalares.

Proposición 5.14 *Si* $\mathbf{f} : \Omega \to F$ *es diferenciable en* $\mathbf{a} \in \Omega$ *se verifica*

a) $\mathbf{f}$ *es continua en* $\mathbf{a}$.

b) Para cada $\mathbf{v} \in E$ *existe la derivada* $D_{\mathbf{v}} f(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{v}$.

DEM:
a) Como $L = d\mathbf{f}(\mathbf{a})$ es una aplicación lineal continua la diferencia $\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = L(\mathbf{h}) + \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h})$ tiende hacia $\mathbf{0}$ cuando $\mathbf{h} \to \mathbf{0}$.

b) Para cada $\mathbf{v} \in E$ se tiene

$$\frac{\mathbf{f}(\mathbf{a}+t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = \frac{L(t\mathbf{v})}{t} + \frac{\|t\mathbf{v}\|}{t}\boldsymbol{\rho}(t\mathbf{v}) = L(\mathbf{v}) + \frac{|t|}{t}\|\mathbf{v}\| \boldsymbol{\rho}(t\mathbf{v})$$

Como $\lim_{t \to 0} \boldsymbol{\rho}(t\mathbf{v}) = 0$ se concluye que existe el límite

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{t \to 0} \frac{\mathbf{f}(\mathbf{a}+t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = L(\mathbf{v})$$

■

Cuando $E = \mathbb{R}^n$, dotado de cualquier norma (recuérdese que en $\mathbb{R}^n$ todas las normas son equivalentes) si $\mathbf{f} : \Omega \to F$ es diferenciable en $\mathbf{a} \in \Omega$, para cada vector $\mathbf{v} = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$, en virtud de la linealidad de la diferencial $d\mathbf{f}(\mathbf{a})$ y de la proposición 5.14 el cálculo de la derivada $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{v}$ se reduce al de las derivadas parciales $D_j\mathbf{f}(\mathbf{a})$ ya que

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\left(\sum_{j=1}^{n} v_j\mathbf{e}_j\right) = \sum_{j=1}^{n} v_j d\mathbf{f}(\mathbf{a})\mathbf{e}_j = \sum_{j=1}^{n} v_j D_j\mathbf{f}(\mathbf{a})$$

OBSERVACIÓN: Si $\mathbf{f}$ diferenciable en $\mathbf{a}$, según 5.14, $\mathbf{f}$ es continua en $\mathbf{a}$, para cada $\mathbf{v} \in E$ existe la derivada $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ y la aplicación $\mathbf{v} \to D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$, es lineal continua. Es natural preguntarse si el recíproco es cierto, es decir si toda función $\mathbf{f} : \Omega \to F$ continua en $\mathbf{a}$ que tenga esta propiedad es diferenciable en $\mathbf{a}$. La respuesta negativa la proporciona la función $f : \mathbb{R}^2 \to \mathbb{R}$ del siguiente ejemplo.

Ejemplo 5.15 *La función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$f(x, y) = \frac{xy^3}{x^2 + y^4} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

es continua en $(0, 0)$, para cada $\mathbf{v} \in \mathbb{R}^2$ existe la derivada $D_{\mathbf{v}}f(0, 0) = 0$, y por lo tanto la aplicación $\mathbf{v} \to D_{\mathbf{v}}f(0, 0)$ es lineal, pero f no es diferenciable en $(0, 0)$.

SOLUCIÓN

Aplicando la desigualdad $|ab| \leq a^2 + b^2$ con $a = x$, $b = y^2$ se obtiene

$$\left|\frac{xy^3}{x^2 + y^4}\right| \leq |y|$$

de donde se sigue que f es continua en $(0, 0)$.

Por otra parte, si $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ se verifica

$$\frac{f(tv_1, tv_2) - f(0, 0)}{t} = \frac{tv_1v_2^3}{v_1^2 + t^2v_2^4}$$

y pasando al límite cuando $t \to 0$ se obtiene que $D_{\mathbf{v}}f(0, 0) = 0$ para todo $\mathbf{v} \in \mathbb{R}^2$

Para ver que f no es diferenciable en $\mathbf{a} = (0, 0)$ podemos razonar por reducción al absurdo: Si lo fuese, en virtud de la proposición 5.14, su diferencial debería ser la aplicación lineal nula, $L(\mathbf{h}) = 0$ para todo $\mathbf{h} \in \mathbb{R}^2$, pero esta aplicación lineal no cumple la condición de diferenciabilidad: El cociente

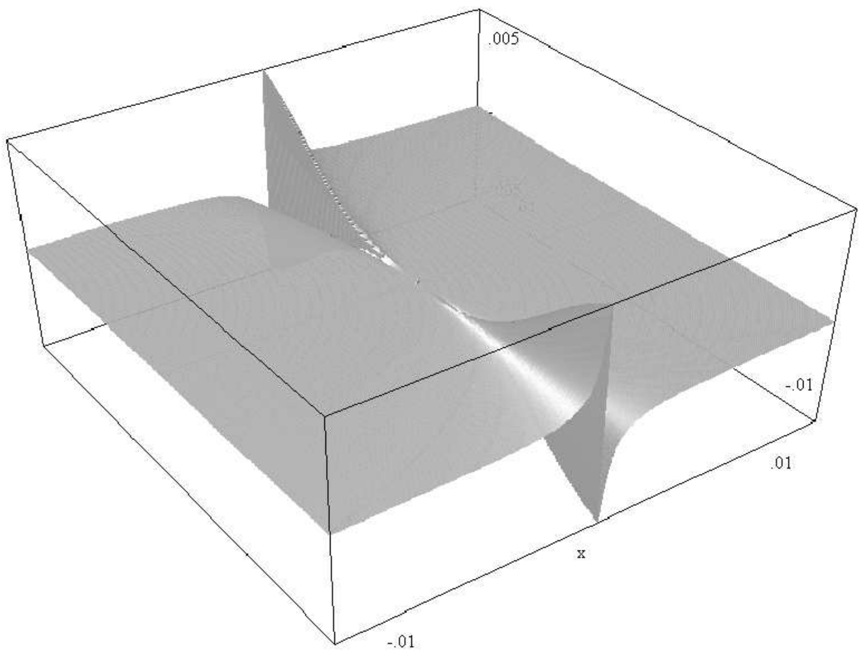
$$\frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = \frac{\mathbf{f}(\mathbf{h})}{\|\mathbf{h}\|} = \frac{h_1h_2^3}{(h_1^2 + h_2^4)\sqrt{h_1^2 + h_2^2}}$$

no tiende hacia 0 cuando $(h_1, h_2) \to (0, 0)$ ya que a través de la parábola $h_2 = t$, $h_1 = t^2$ se tiene

$$\frac{\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a})}{\|\mathbf{h}\|} = \frac{t^5}{2t^4\sqrt{t^4 + t^2}} = \frac{t}{2|t|\sqrt{1 + t^2}}$$

que tiende hacia 1/2 (resp. −1/2 cuando $t \to 0+$ (resp. $t \to 0-$). ■

La función que interviene en 5.15 proporciona un ejemplo interesante de una función continua tal que los vectores $\{(\mathbf{v}, D_{\mathbf{v}}f(0,0)) : \mathbf{v} \in \mathbb{R}^2\}$ forman un plano, candidato a ser plano tangente, y sin embargo f no es diferenciable en $(0,0)$. Para ver gráficamente la razón de la no diferenciabilidad de f en $(0,0)$ se recomienda utilizar un programa de ordenador como DpGraph que permita visualizar la gráfica de f en un entorno de $(0,0)$.



Podremos apreciar que por muy potente que sea el zoom que apliquemos en un entorno de $(0,0,0)$ siempre veremos un pliegue a lo largo del eje Oy y nunca lograremos que la gráfica de f se confunda con el plano $z = 0$. Esto se debe a que en este ejemplo la aplicación $\mathbf{v} \to D_{\mathbf{v}}f(0,0)$, aunque es lineal no proporciona una aproximación local del incremento $f(v_1, v_2) - f(0,0)$ que sea uniforme en todas las direcciones. El ejercicio resuelto 5.42 tiene como finalidad aclarar lo que acabamos de comentar en relación con la función del ejemplo 5.15

Condición suficiente de diferenciabilidad. Si una función de varias variables reales es diferenciable en un punto $\mathbf{a}$ sabemos que su diferencial tiene que ser la aplicación lineal

$$\mathbf{v} \to d\mathbf{f}(\mathbf{a})\mathbf{v} = D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \sum_{j=1}^{n} v_j D_j \mathbf{f}(\mathbf{a})$$

Esto nos indica la estrategia natural para estudiar la diferenciabilidad, en un punto $\mathbf{a}$, de una función $\mathbf{f}$ de varias variables reales: Se comienza estudiando la existencia de las derivadas parciales $D_j\mathbf{f}(\mathbf{a})$, $1 \leq j \leq n$, y en el caso de que existan todas se considera el candidato natural para la diferencial $d\mathbf{f}(\mathbf{a})$, que es la aplicación lineal

$L : \mathbb{R}^n \to F$ definida por $L(\mathbf{u}) = \sum_{j=1}^n u_j D_j \mathbf{f}(\mathbf{a})$. Finalmente se estudia si el cociente $\boldsymbol{\rho}(\mathbf{h}) = [\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})] / \|\mathbf{h}\|$ tiende hacia $\mathbf{0}$ cuando $\mathbf{h} \to \mathbf{0}$.

Esta estrategia es la que se usa para demostrar el siguiente teorema que proporciona una condición suficiente de diferenciabilidad, muy útil en la práctica. En las hipótesis parece que desempeña un papel especial la última variable, pero su papel lo puede desempeñar cualquier otra

Teorema 5.16 [Condición suficiente de diferenciabilidad] *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$ *con valores en un espacio normado* $(F, \| \ \|)$. *Se supone que existe la derivada parcial* $D_n \mathbf{f}(\mathbf{a})$ *y que en una bola* $B(\mathbf{a}, r) \subset \Omega$ *existen las otras derivadas parciales* $D_1\mathbf{f}(\mathbf{x}), D_2\mathbf{f}(\mathbf{x}), \cdots D_{n-1}\mathbf{f}(\mathbf{x})$, *y son continuas en* $\mathbf{a}$. *Entonces* $\mathbf{f}$ *es diferenciable en* $\mathbf{a}$.

DEM: Para simplificar la exposición usamos en $\mathbb{R}^n$ la norma $\| \ \|_\infty$, de modo que

$$B(\mathbf{a}, r) = \{(x_1, x_2, \cdots, x_n) : |x_j - a_j| < r, 1 \le j \le n\}$$

Como las funciones $D_j\mathbf{f}(\mathbf{x})$, $1 \le j < n$ son continuas en $\mathbf{a}$, dado $\epsilon > 0$, existe $0 < \delta < r$ tal que

$$(*) \qquad \|\mathbf{y} - \mathbf{a}\|_\infty < \delta \Rightarrow \|D_j\mathbf{f}(\mathbf{y}) - D_j\mathbf{f}(\mathbf{a})\| < \epsilon \text{ para } 1 \le j < n.$$

Para cada $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in B(\mathbf{a}, \delta)$, la función de variable real

$$\boldsymbol{\varphi}_1(t) = \mathbf{f}(t, x_2, \cdots, x_n) - (t - a_1) D_1\mathbf{f}(a_1, a_2, \cdots, a_n)$$

está definida y es derivable para $|t - a_1| < \delta$. En virtud de (*), su derivada

$$\boldsymbol{\varphi}_1'(t) = D_1\mathbf{f}(t, x_2, x_3, \cdots, x_n) - D_1\mathbf{f}(a_1, a_2, \cdots, a_n)$$

verifica

$$\|\boldsymbol{\varphi}_1'(t)\| < \epsilon \quad \text{si} \quad |t - a_1| < \delta.$$

Aplicando el teorema del incremento finito a la función $\boldsymbol{\varphi}_1$ en el intervalo $[a_1, x_1]$ se obtiene $\|\boldsymbol{\varphi}_1(x_1) - \boldsymbol{\varphi}_1(a_1)\| \le \epsilon |x_1 - a_1|$ es decir

$$(1) \qquad \|\mathbf{f}(x_1, x_2, \cdots, x_n) - \mathbf{f}(a_1, x_2, \cdots, x_n) - (x_1 - a_1) D_1\mathbf{f}(\mathbf{a})\| \le \epsilon |x_1 - a_1|$$

Razonando de forma similar con

$$\boldsymbol{\varphi}_2(t) = \mathbf{f}(a_1, t, x_3, \cdots, x_n) - (t - a_2) D_2\mathbf{f}(a_1, a_2, \cdots, a_n)$$

se obtiene la desigualdad

$$(2) \qquad \|\mathbf{f}(a_1, x_2, x_3, \cdots, x_n) - \mathbf{f}(a_1, a_2, x_3, \cdots, x_n) - (x_2 - a_2) D_2\mathbf{f}(\mathbf{a})\| \le \epsilon |x_2 - a_2|$$

Análogamente se razona con las restantes variables, hasta la variable x_{n-1}

$$\text{(n-1)} \qquad \|\mathbf{f}(a_1, \cdots, a_{n-2}, x_{n-1}, x_n) - \mathbf{f}(a_1, \cdots, a_{n-1}, x_n) - (x_{n-1} - a_{n-1}) D_{n-1}\mathbf{f}(\mathbf{a})\| \le$$
$$\le \epsilon |x_{n-1} - a_{n-1}|$$

Por último, usando la definición de $D_n\mathbf{f}(\mathbf{a})$ como límite, podemos suponer que $\delta > 0$ se ha elegido de modo que también se cumple

$$\left\| \frac{\mathbf{f}(a_1, \cdots, a_{n-1}, x_n) - \mathbf{f}(a_1, \cdots, a_n)}{x_n - a_n} - D_n\mathbf{f}(\mathbf{a}) \right\| < \epsilon$$

es decir

$$\text{(n)} \qquad \|\mathbf{f}(a_1, \cdots, a_{n-1}, x_n) - \mathbf{f}(a_1, \cdots, a_n) - (x_n - a_n)D_n\mathbf{f}(\mathbf{a})\| < \epsilon|x_n - a_n|$$

Considerando la suma telescópica

$$\begin{array}{lr}
\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \sum_{i=j}^{n}(x_j - a_j)D_j\mathbf{f}(\mathbf{a}) = & \\
\mathbf{f}(x_1, x_2, \cdots, x_n) - \mathbf{f}(a_1, x_2, \cdots, x_n) - (x_1 - a_1)D_1\mathbf{f}(\mathbf{a}) & + \\
\mathbf{f}(a_1, x_2, \cdots, x_n) - \mathbf{f}(a_1, a_2, \cdots, x_n) - (x_2 - a_2)D_2\mathbf{f}(\mathbf{a}) & + \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & + \\
\mathbf{f}(a_1, \cdots a_{n-1}, x_n) - \mathbf{f}(a_1, a_2 \cdots, a_n) - (x_n - a_n)D_n\mathbf{f}(\mathbf{a}) &
\end{array}$$

y combinando las desigualdades (1), (2), (n-1), (n), con la desigualdad triangular se obtiene que cuando $\|\mathbf{x} - \mathbf{a}\|_\infty < \delta$ se cumple

$$\left\| \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - \sum_{j=1}^{n}(x_j - a_j)D_j\mathbf{f}(\mathbf{a}) \right\| \leq \epsilon \sum_{j=1}^{n} |x_j - a_j| \leq n\epsilon \|\mathbf{x} - \mathbf{a}\|_\infty$$

Con esto queda demostrado que $\mathbf{f}$ es diferenciable en $\mathbf{a}$ y que su diferencial es la aplicación lineal $L : \mathbb{R}^n \to F$ dada por $L(\mathbf{h}) = \sum_{j=1}^{n} h_j D_j\mathbf{f}(\mathbf{a})$ ■

Definición 5.17 *Si $\Omega \subset \mathbb{R}^n$ es abierto y $\mathbf{f} : \Omega \to \mathbf{F}$ tiene derivadas parciales continuas en todo $\mathbf{x} \in \Omega$ se dice que $\mathbf{f}$ es de clase C^1 en Ω y se escribe $\mathbf{f} \in C^1(\Omega, F)$. A las aplicaciones de clase C^1 también se les llama continuamente diferenciables.*

En virtud del teorema 5.16 cada aplicación de clase C^1 es diferenciable en todos los puntos de su dominio.

5.3. Las reglas del cálculo diferencial

En esta sección se exponen los resultados típicos sobre diferenciabilidad de funciones que se obtienen combinando otras funciones diferenciables mediante las operaciones usuales (suma, producto, cociente, composición, etc). Al mismo tiempo obtendremos las reglas para obtener su diferencial en términos de las diferenciales de las funciones que intervienen en su definición. El resultado más notable es la regla de la cadena de la que se desprenden las reglas para el cálculo de derivadas parciales de funciones compuestas. La regla de la cadena también servirá para demostrar que cuando $\mathbf{f}$ es una función de k variables reales diferenciable en $\mathbf{a}$ entonces el conjunto de los vectores tangentes a su gráfica en el punto $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$ forma un espacio vectorial de dimensión k. Comenzamos viendo que para funciones con valores en $\mathbb{R}^m$ el estudio de su diferenciabilidad se puede realizar componente a componente:

Proposición 5.18 *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^m$ *de componentes* $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots f_m(\mathbf{x}))$, *definida en un abierto* $\Omega \subset E$. *Entonces* $\mathbf{f}$ *es diferenciable en* $\mathbf{a} \in \Omega$ *si y sólo si todas sus componentes lo son y en este caso* $L = d\mathbf{f}(\mathbf{a}) : E \to \mathbb{R}^m$ *es la aplicación lineal continua cuyas componentes* $L = (L_1, L_2, \cdots L_m)$ *son las diferenciales de las componentes de* $\mathbf{f}$, *es decir* $df_j(\mathbf{a}) = L_j$, $j = 1, 2 \cdots m$.

Dem: Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ su diferencial $L = d\mathbf{f}(\mathbf{a})$ verifica

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \to 0} \boldsymbol{\rho}(\mathbf{h}) = 0$$

y pasando a las componentes resulta que para cada $j \in \{1, 2, \cdots m\}$ se cumple

$$f_j(\mathbf{a}+\mathbf{h}) - f_j(\mathbf{a}) - L_j(\mathbf{h}) = \|\mathbf{h}\| \rho_j(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \to 0} \rho_j(\mathbf{h}) = 0$$

luego cada f_j es diferenciable en $\mathbf{a}$ y $df_j(\mathbf{a}) = L_j$.

Recíprocamente, si cada f_j es diferenciable en $\mathbf{a}$ y $L_j = df_j(\mathbf{a})$, entonces la aplicación lineal continua $L : E \to F$ formada con estas componentes, $L = (L_1, L_2, \cdots L_m)$ cumple

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}(\mathbf{h})$$

donde $\boldsymbol{\rho}(\mathbf{h}) = (\rho_1(\mathbf{h}), \rho_2(\mathbf{h}), \cdots \rho_m(\mathbf{h}))$ tiende hacia 0 cuando $\mathbf{h} \to 0$. Por lo tanto $\mathbf{f}$ es diferenciable en $\mathbf{a}$ y $d\mathbf{f}(\mathbf{a}) = L$. ■

Proposición 5.19 *Si* $\mathbf{f}, \mathbf{g} : \Omega \to F$ *son diferenciables en* $\mathbf{a} \in \Omega$, *donde* $\Omega \subset E$ *es abierto, se verifica:*

i) $\mathbf{f} + \mathbf{g}$ *es diferenciable en* $\mathbf{a}$ *y* $d(\mathbf{f} + \mathbf{g})(\mathbf{a}) = d\mathbf{f}(\mathbf{a}) + d\mathbf{g}(\mathbf{a})$.

ii) Si la norma de F *procede del producto escalar* $\langle \mid \rangle$, *la función*

$$\psi(\mathbf{x}) = \langle \mathbf{f}(\mathbf{x}) \mid \mathbf{g}(\mathbf{x}) \rangle$$

es diferenciable en $\mathbf{a}$ *y su diferencial* $d\psi(\mathbf{a}) : E \to \mathbb{R}$ *es la aplicación lineal*

$$d\psi(\mathbf{a})\mathbf{v} = \langle \mathbf{f}(\mathbf{a}) \mid d\mathbf{g}(\mathbf{a})\mathbf{v} \rangle + \langle d\mathbf{f}(\mathbf{a})\mathbf{v} \mid \mathbf{g}(\mathbf{a}) \rangle$$

iii) Si $\alpha : \Omega \to \mathbb{R}$ *es diferenciables en* $\mathbf{a} \in \Omega$ *entonces* $\mathbf{g}(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{f}(\mathbf{x})$ *es diferenciable en* $\mathbf{a}$ *y* $d(\mathbf{g})(\mathbf{a}) : E \to E$ *es la aplicación lineal dada por*

$$d\mathbf{g}(\mathbf{a})\mathbf{v} = \alpha(\mathbf{a})d\mathbf{f}(\mathbf{a})\mathbf{v} + [d\alpha(\mathbf{a})\mathbf{v}]\mathbf{f}(\mathbf{a})$$

iv) Si $\alpha : \Omega \to \mathbb{R}$ *es diferenciables en* $\mathbf{a} \in \Omega$ *y* $\alpha(\mathbf{a}) \neq 0$, *entonces la función* $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{x})/\alpha(\mathbf{x})$, *(definida en un entorno de* $\mathbf{a}$*) es diferenciable en* $\mathbf{a}$ *y*

$$d(\mathbf{g})(\mathbf{a})\mathbf{v} = \frac{\alpha(\mathbf{a})d\mathbf{f}(a)\mathbf{v} - [d\alpha(\mathbf{a})\mathbf{v}]\mathbf{f}(\mathbf{a})}{\alpha(\mathbf{a})^2}$$

Dem: La demostración de i) es inmediata. Para demostrar ii) veamos primero la continuidad de la aplicación lineal $S : E \to \mathbb{R}$, definida por

$$S(\mathbf{v}) = \langle \mathbf{f}(\mathbf{a}) \mid T(\mathbf{v})\rangle + \langle L(\mathbf{v}) \mid \mathbf{g}(\mathbf{a})\rangle \text{ donde } L = d\mathbf{f}(\mathbf{a}) \text{ y } T = d\mathbf{g}(\mathbf{a})$$

En virtud de la desigualdad de Cauchy-Schwarz

$$|S(\mathbf{v})| \leq \|\mathbf{f}(\mathbf{a})\| \, \|T(\mathbf{v})\| + \|L(\mathbf{v})\| \, \|\mathbf{g}(\mathbf{a})\| \leq M \|\mathbf{v}\|$$

con $M = \|\mathbf{f}(\mathbf{a})\| \, \|T\| + \|\mathbf{g}(\mathbf{a})\| \, \|L\|$, y por lo tanto S es una aplicación lineal continua. (Cuando E es finito dimensional no es preciso demostrar esto). Por hipótesis,

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) + \boldsymbol{\alpha}(\mathbf{h}), \quad \mathbf{g}(\mathbf{a}+\mathbf{h}) = \mathbf{g}(\mathbf{a}) + T(\mathbf{h}) + \boldsymbol{\beta}(\mathbf{h})$$

donde $\boldsymbol{\alpha}(\mathbf{h}) = o(\|\mathbf{h}\|)$ y $\boldsymbol{\beta}(\mathbf{h}) = o(\|\mathbf{h}\|)$. Por la bilinealidad del producto escalar

$$\langle \mathbf{f}(\mathbf{a}+\mathbf{h}) \mid \mathbf{g}(\mathbf{a}+\mathbf{h})\rangle = \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}+\mathbf{h})\rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}+\mathbf{h})\rangle =$$

$$= \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}) + T(\mathbf{h})\rangle + \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \boldsymbol{\beta}(\mathbf{h})\rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}+\mathbf{h})\rangle =$$
$$= \langle \mathbf{f}(\mathbf{a}) \mid \mathbf{g}(\mathbf{a})\rangle + S(\mathbf{h}) + \boldsymbol{\rho}(\mathbf{h})$$

donde

$$\boldsymbol{\rho}(\mathbf{h}) = \langle L(\mathbf{h}) \mid T(\mathbf{h})\rangle + \langle \mathbf{f}(\mathbf{a}) + L(\mathbf{h}) \mid \boldsymbol{\beta}(\mathbf{h})\rangle + \langle \boldsymbol{\alpha}(\mathbf{h}) \mid \mathbf{g}(\mathbf{a}+\mathbf{h})\rangle$$

y basta ver que $\boldsymbol{\rho}(\mathbf{h}) = o(\|\mathbf{h}\|)$: En virtud de desigualdad de Cauchy-Schwarz

$$|\boldsymbol{\rho}(\mathbf{h})| \leq \|L\| \, \|T\| \, \|\mathbf{h}\|^2 + \|\mathbf{f}(\mathbf{a}) + L(\mathbf{h})\| \, \|\boldsymbol{\beta}(\mathbf{h})\| + \|\mathbf{g}(\mathbf{a}+\mathbf{h})\| \, \|\boldsymbol{\alpha}(\mathbf{h})\|$$

Después de dividir por $\|\mathbf{h}\|$, los tres términos de la derecha tienden hacia 0 cuando $\mathbf{h} \to \mathbf{0}$, luego $|\boldsymbol{\rho}(\mathbf{h})| = o(\|\mathbf{h}\|)$.

La demostración de iii) que es análoga a la de ii), se deja al cuidado del lector, y iv) es consecuencia directa de iii). ∎

Teorema 5.20 [Regla de la cadena] *Sean $(E, \| \ \|)$, $(F, \| \ \|)$ y $(G, \| \ \|)$ espacios normados y $\Omega \subset E$, $V \subset F$ abiertos. Si $\mathbf{f} : \Omega \to V$ es diferenciable en $\mathbf{a} \in \Omega$ y $\mathbf{g} : V \to G$ es diferenciable en $\mathbf{b} = \mathbf{f}(\mathbf{a})$, entonces $\mathbf{g} \circ \mathbf{f}$ es diferenciable en $\mathbf{a}$ y*

$$d(\mathbf{g} \circ \mathbf{f})(\mathbf{a}) = d\mathbf{g}(\mathbf{f}(\mathbf{a})) \circ d\mathbf{g}(\mathbf{a})$$

Dem: Si $L = d\mathbf{f}(\mathbf{a})$ y $T = d\mathbf{g}(\mathbf{b})$, en virtud de la hipótesis:

i) $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + L(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\| \, \boldsymbol{\rho}_1(\mathbf{x})$ con $\lim_{\mathbf{x} \to \mathbf{a}} \boldsymbol{\rho}_1(\mathbf{x}) = 0$.

ii) $\mathbf{g}(\mathbf{y}) = \mathbf{g}(\mathbf{b}) + T(\mathbf{y} - \mathbf{b}) + \|\mathbf{y} - \mathbf{b}\| \, \boldsymbol{\rho}_2(\mathbf{y})$, con $\lim_{\mathbf{y} \to \mathbf{b}} \boldsymbol{\rho}_2(\mathbf{y}) = 0$

En lo que sigue conviene suponer que hemos definido $\rho_1(\mathbf{a}) = \mathbf{0} \in F$ y $\rho_2(\mathbf{b}) = \mathbf{0} \in G$. Así la función $\rho_1(\mathbf{x})$ (resp. $\rho_2(\mathbf{y})$) está definida para todo $\mathbf{x} \in \Omega$ (resp. $\mathbf{y} \in V$) y es continua en el punto $\mathbf{x} = \mathbf{a}$ (resp. $\mathbf{y} = \mathbf{b}$). Sustituyendo $\mathbf{y} = \mathbf{f}(\mathbf{x})$ en ii) resulta

$$\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{g}(\mathbf{f}(\mathbf{a})) + T(\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})) + \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \,\boldsymbol{\rho}_2(\mathbf{f}(\mathbf{x}))$$

y reemplazando $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})$ por el valor que proporciona i)

$$(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = (\mathbf{g} \circ \mathbf{f})(\mathbf{a}) + (T \circ L)(\mathbf{x} - \mathbf{a}) + \|\mathbf{x} - \mathbf{a}\| \,\mathbf{r}(\mathbf{x})$$

donde

$$\mathbf{r}(\mathbf{x}) = T(\boldsymbol{\rho}_1(\mathbf{x})) + \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|}\boldsymbol{\rho}_2(\mathbf{f}(\mathbf{x}))$$

y para terminar basta demostrar que $\lim_{\mathbf{x} \to \mathbf{a}} \mathbf{r}(\mathbf{x}) = 0$.

Cuando $\mathbf{x} \to \mathbf{a}$, en virtud de la continuidad de T y de $\mathbf{f}$, podemos asegurar que $T(\boldsymbol{\rho}_1(\mathbf{x}))$ tiende hacia 0 y que $\mathbf{f}(\mathbf{x})$ tiende hacia $\mathbf{b} = \mathbf{f}(\mathbf{a})$, y teniendo en cuenta que ρ_2 es continua en $\mathbf{b}$ se obtiene que $\lim_{\mathbf{x} \to \mathbf{a}} \boldsymbol{\rho}_2(\mathbf{f}(\mathbf{x})) = \rho_2(\mathbf{b}) = \mathbf{0}$. Por consiguiente, para demostrar lo que se desea, basta ver que el cociente

$$\frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|}{\|\mathbf{x} - \mathbf{a}\|}$$

se mantiene acotado en un entorno de $\mathbf{a}$. Como $\mathbf{f}$ es diferenciable en $\mathbf{a}$, existe $\delta > 0$ tal que si $\|\mathbf{x} - \mathbf{a}\| < \delta$ se cumple $\|\boldsymbol{\rho}_1(\mathbf{x})\| < 1$ luego, en virtud de i)

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| \leq \|L\| \,\|\mathbf{x} - \mathbf{a}\| + \|\mathbf{x} - \mathbf{a}\| \,\|\boldsymbol{\rho}_1(\mathbf{x})\| \leq (\|L\| + 1) \,\|\mathbf{x} - \mathbf{a}\|$$

■

Corolario 5.21 *Sean $(F, \| \|)$, $(G, \| \|)$ espacios normados, y $\Omega \subset \mathbb{R}$, $V \subset F$ abiertos. Si $\mathbf{f} : \Omega \to V$ es derivable en $t \in \Omega$ y $\mathbf{g} : V \to G$ diferenciable en $\mathbf{f}(t)$, entonces $\mathbf{g} \circ \mathbf{f}$ es derivable en t y*

$$(\mathbf{g} \circ \mathbf{f})'(t) = d\mathbf{g}(\mathbf{f}(t))\mathbf{f}'(t)$$

es decir el vector $(\mathbf{g} \circ \mathbf{f})'(t) \in G$ es la imagen del vector $\mathbf{f}'(t) \in F$ mediante la aplicación lineal $d\mathbf{g}(\mathbf{f}(t)) : F \to G$.

Dem: Es consecuencia directa de 5.20 pues para cada $h \in \mathbb{R}$ se tiene

$$h(\mathbf{g} \circ \mathbf{f})'(t) = d(\mathbf{g} \circ \mathbf{f})(t)h = d\mathbf{g}(\mathbf{f}(t))(d\mathbf{f}(t)h) = d\mathbf{g}(\mathbf{f}(t))(h\mathbf{f}'(t)) = hd\mathbf{g}(\mathbf{f}(t))\mathbf{f}'(t)$$

■

Expresión analítica de la diferencial. Matriz Jacobiana. Supongamos que $\Omega \subset \mathbb{R}^n$ es abierto y que $\mathbf{f} : \Omega \to \mathbb{R}^m$ es diferenciable en $\mathbf{a} \in \Omega$. Su diferencial es una aplicación lineal $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$ que se suele identificar con una matriz de m filas y n columnas. Para obtener la matriz de $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$ respecto a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$, basta considerar las componentes $(f_1, f_2, \cdots f_m)$ de $\mathbf{f}$. Según

5.18 cada componente f_j es diferenciable en $\mathbf{a}$ y $df_j(\mathbf{a}) = L_j$ donde $(L_1, L_2, \cdots L_m)$ son las componentes de $L = d\mathbf{f}(\mathbf{a})$. Entonces, para cada $\mathbf{v} = (v_1, v_2, \cdots v_n) \in \mathbb{R}^n$, su imagen $d\mathbf{f}(\mathbf{a})\mathbf{v} = \mathbf{u} = (u_1, u_2, \cdots u_m)$ viene dada por

$$u_j = L_j(\mathbf{v}) = df_j(a)\mathbf{v} = \sum_{k=1}^{m} D_k f_j(\mathbf{a}) v_k = \sum_{k=1}^{m} \frac{\partial f_j(\mathbf{a})}{\partial x_k} v_k$$

fórmula que con notación matricial se escribe en la forma

$$\begin{pmatrix} u_1 \\ u_2 \\ \cdots \\ u_m \end{pmatrix} = \begin{pmatrix} D_1 f_1(\mathbf{a}) & D_2 f_1(\mathbf{a}) & \cdots & D_n f_1(\mathbf{a}) \\ D_1 f_2(\mathbf{a}) & D_2 f_2(\mathbf{a}) & \cdots & D_n f_2(\mathbf{a}) \\ \cdots & \cdots & \cdots & \cdots \\ D_1 f_m(\mathbf{a}) & D_2 f_m(\mathbf{a}) & \cdots & D_n f_m(\mathbf{a}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{pmatrix}$$

La matriz anterior es la matriz de la diferencial $df(\mathbf{a})$ respecto a las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. Se le llama matriz jacobiana y también matriz derivada. Utilizando la notación $\mathbf{f}'(\mathbf{a})$ para designar esta matriz podemos escribir $d\mathbf{f}(\mathbf{a})\mathbf{v} = \mathbf{f}'(\mathbf{a})\mathbf{v}$, con el convenio de que $\mathbf{f}'(\mathbf{a})\mathbf{v}$ es el producto de la matriz $\mathbf{f}'(\mathbf{a})$ por el vector $\mathbf{v}$ escrito en forma de vector columna. Notaciones más habituales para la matriz jacobiana son

$$\frac{D(f_1, f_2, \cdots f_m)}{D(x_1, x_2, \cdots x_n)}(\mathbf{a}), \quad \frac{\partial(f_1, f_2, \cdots f_m)}{\partial(x_1, x_2, \cdots x_n)}(\mathbf{a})$$

En particular, cuando $n = 1$, sabemos que $\mathbf{f}$ es derivable en a si y sólo si es diferenciable en a y en ese caso su diferencial es la aplicación lineal $d\mathbf{f}(a) : \mathbb{R} \to \mathbb{R}^m$, definida por $d\mathbf{f}(a)(h) = \mathbf{f}'(a)h$ y la matriz de la diferencial $d\mathbf{f}(a)$ se identifica con $\mathbf{f}'(a)$ escrito como vector columna.

Consideremos ahora el caso de una función compuesta $\mathbf{g} \circ \mathbf{f}$ donde $\mathbf{f} : \Omega \to V$ está definida en un abierto $\Omega \subset \mathbb{R}^n$ y $\mathbf{g} : V \to \mathbb{R}^p$ en un abierto $V \subset \mathbb{R}^m$. Se supone que $\mathbf{f}$ es diferenciable en $\mathbf{x} \in \Omega$ y que $\mathbf{g}$ diferenciable en $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Entonces, de acuerdo con 5.20, la matriz jacobiana de la función compuesta $\boldsymbol{\varphi} = \mathbf{g} \circ \mathbf{f}$, en el punto $\mathbf{x}$, es el producto de la matrices jacobianas, $\boldsymbol{\varphi}'(\mathbf{x}) = \mathbf{g}'(\mathbf{y})\mathbf{f}'(\mathbf{x})$, es decir

$$\begin{pmatrix} D_1\varphi_1(\mathbf{x}) & \cdots & D_n\varphi_1(\mathbf{x}) \\ D_1\varphi_2(\mathbf{x}) & \cdots & D_n\varphi_2(\mathbf{x}) \\ \cdots & \cdots & \cdots \\ D_1\varphi_p(\mathbf{x}) & \cdots & D_n\varphi_p(\mathbf{x}) \end{pmatrix} =$$

$$= \begin{pmatrix} D_1 g_1(\mathbf{y}) & \cdots & D_m g_1(\mathbf{y}) \\ D_1 g_2(\mathbf{y}) & \cdots & D_m g_2(\mathbf{y}) \\ \cdots & \cdots & \cdots \\ D_1 g_p(\mathbf{y}) & \cdots & D_m g_p(\mathbf{y}) \end{pmatrix} \begin{pmatrix} D_1 f_1(\mathbf{x}) & \cdots & D_n f_1(\mathbf{x}) \\ D_1 f_2(\mathbf{x}) & \cdots & D_n f_2(\mathbf{x}) \\ \cdots & \cdots & \cdots \\ D_1 f_m(\mathbf{x}) & \cdots & D_n f_m(\mathbf{x}) \end{pmatrix}$$

y de acuerdo con la fórmula para el producto de matrices:

$$D_i\boldsymbol{\varphi}_j(\mathbf{x}) = \sum_{k=1}^{m} D_k \mathbf{g}_j(\mathbf{f}(\mathbf{x})) D_i \mathbf{f}_k(\mathbf{a}), \quad i = 1, \cdots, n, \ j = 1, 2, \cdots, p \tag{5.1}$$

Con el fin de recordar fácilmente esta regla para el cálculo de las derivadas parciales de una función compuesta es conveniente adoptar el siguiente convenio: Denotemos por $(x_1, x_2, \cdots x_n)$ las variables independientes de $\mathbf{f}$, y por $(y_1, y_2, \cdots y_n)$ las variables independientes de $\mathbf{g}$, de modo que para obtener la función compuesta

$$(z_1, z_2, \cdots z_p) = (\varphi_1(x_1, x_2, \cdots x_n), \varphi_2(x_1, x_2, \cdots x_n), \cdots \varphi_p(x_1, x_2, \cdots x_n))$$

basta sustituir $y_k = f_k(x_1, x_2, \cdots x_n)$, $1 \leq k \leq m$, en la expresión

$$(z_1, z_2, \cdots z_p) = \mathbf{g}(y_1, y_2, \cdots y_m)$$

Si hacemos el convenio de designar del mismo modo a las variables y_k y a las funciones f_k (y a las variables z_j y a las funciones g_j) la fórmula para las derivadas parciales de la función compuesta se escribe en una forma fácil de recordar.

$$\frac{\partial z_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k} \frac{\partial y_k}{\partial x_i}$$

donde, con el fin de simplificar la expresión, hemos omitido los puntos donde se evalúan las derivadas parciales. Esto es lo que se hace habitualmente cuando los puntos quedan claros por el contexto.

Las notaciones del cálculo diferencial Consideremos primero el caso de una función real de n variables reales que es diferenciable en $\mathbf{a} \in \Omega \subset \mathbb{R}^n$. Su diferencial $df(\mathbf{a})$ pertenece al espacio vectorial n-dimensional $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ y una base de este espacio vectorial son las proyecciones

$$\pi_k : \mathbb{R}^n \to \mathbb{R}; \quad \pi_k(x_1, x_2, \cdots x_n) = x_k$$

En cálculo diferencial la proyección π_k se acostumbra a denotar dx_k. La razón de esto es la siguiente: Resulta cómodo denotar una función por la fórmula que se usa para definirla. Así, por ejemplo, la función $f : \mathbb{R}^3 \to \mathbb{R}$ que asigna al vector $\mathbf{x} = (x, y, z)$ el número $x \cos(y + e^z)$ se suele designar mediante la fórmula que la define. Con este convenio, si estamos trabajando con funciones de tres variables, la función x será la proyección $\pi_1 : \mathbb{R}^3 \to \mathbb{R}$ que asocia al vector $(x, y, z) \in \mathbb{R}^3$ su primera coordenada x. Esta proyección es lineal y por lo tanto su diferencial dx es la aplicación constante que a cada punto $\mathbf{a} \in \mathbb{R}^3$ le hace corresponder la proyección π_1 (véase el ejemplo ii) de la sección 2). Siguiendo el convenio habitual de identificar una aplicación constante con su valor constante, es natural utilizar la notación dx para designar la proyección π_1. Del mismo modo, dy y dz designan las proyecciones π_2 y π_3, respectivamente. Análogamente, en $\mathbb{R}^n$, la proyección π_k se denota dx_k $(k = 1 \cdots n)$

Sea $\{\mathbf{e}_k : 1 \leq k \leq n\}$ la base canónica de $\mathbb{R}^n$. Dado $\mathbf{v} = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$, en virtud de la linealidad de la diferencial $df(\mathbf{a})$ y de la proposición 5.14 sabemos que

$$df(\mathbf{a})\mathbf{v} = \sum_{k=1}^{n} D_k f(\mathbf{a}) v_k$$

Como v_k es la imagen de $\mathbf{v}$ mediante la proyección dx_k resulta

$$df(\mathbf{a})\mathbf{v} = \left(\sum_{k=1}^{n} D_k f(\mathbf{a}) dx_k\right) \mathbf{v}$$

luego las derivadas parciales $(D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \cdots D_n f(\mathbf{a}))$ son las coordenadas de $df(\mathbf{a})$ respecto a la base canónica $(dx_1, dx_2, \cdots, \cdots dx_n)$ de $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$:

$$df(\mathbf{a}) = \sum_{k=1}^{n} D_k f(\mathbf{a}) dx_k = \sum_{k=1}^{n} \frac{\partial f(\mathbf{a})}{\partial x_k} dx_k$$

En particular, cuando $n = 2$ se acostumbra a escribir

$$df(\mathbf{a}) = \frac{\partial f(\mathbf{a})}{\partial x} dx + \frac{\partial f(\mathbf{a})}{\partial y} dy = f_x(\mathbf{a}) dx + f_y(\mathbf{a}) dy$$

Análogamente, cuando $n = 3$

$$df(\mathbf{a}) = \frac{\partial f(\mathbf{a})}{\partial x} dx + \frac{\partial f(\mathbf{a})}{\partial y} dy + \frac{\partial f(\mathbf{a})}{\partial z} dz = f_x(\mathbf{a}) dx + f_y(\mathbf{a}) dy + f_z(\mathbf{a}) dz$$

En el caso de funciones de una variable se tiene $df(a) = f'(a)dx$ lo que está de acuerdo con la notación de Leibniz para la derivada $\mathbf{f}'(a) = \frac{d\mathbf{f}}{dx}(a)$, donde la expresión $\frac{d\mathbf{f}}{dx}$ se debe entender como un símbolo unitario que designa a la derivada $\mathbf{f}'$ y no como una fracción. Este simbolismo permite escribir, omitiendo el punto a, la fórmula clásica de Leibniz

$$d\mathbf{f} = \frac{d\mathbf{f}}{dx} dx$$

que actualmente se interpreta como una igualdad entre las aplicaciones lineales $d\mathbf{f}(a)$ y $\mathbf{f}'(a)dx$. La flexibilidad de la notación de Leibniz se pone de manifiesto con la regla de la cadena 4.6: En términos de diferenciales, la regla de la cadena para la composición $g(t) = f(\varphi(t))$ adopta la forma $dg(a)h = g'(a)h = f'(\varphi(a))\varphi'(a)h = f'(\varphi(a))[d\varphi(a)h] = [d\mathbf{f}(\varphi(a)) \circ d\varphi(a)]h$ es decir $dg(a) = df(\varphi(a)) \circ d\varphi(a)$.

El teorema del incremento finito. En lo que sigue, si E es un espacio vectorial normado el segmento cerrado (resp. abierto) de extremos $\mathbf{a}, \mathbf{b} \in E$, es

$$[\mathbf{a}, \mathbf{b}] = \{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) : 0 \leq t \leq 1\}, \text{ (resp. } (\mathbf{a}, \mathbf{b}) = \{\mathbf{a} + t(\mathbf{b} - \mathbf{a}) : 0 < t < 1\})$$

Teorema 5.22 [Incremento finito] *Sea* $\mathbf{f} : \Omega \to F$*, definida en un abierto* Ω *del espacio normado* $(E, \| \|)$*, con valores en el espacio normado* $(F, \| \|)$*, continua en cada punto del segmento cerrado* $[\mathbf{a}, \mathbf{b}]$ *y diferenciable en cada punto del segmento abierto* $(\mathbf{a}, \mathbf{b})$ *con* $\|d\mathbf{f}(\mathbf{x})\| \leq M$ *para todo* $\mathbf{x} \in (\mathbf{a}, \mathbf{b})$*. Entonces* $\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq M \|\mathbf{b} - \mathbf{a}\|$*.*

Dem: En virtud de 5.21 la función de variable real $\boldsymbol{\varphi}(t) = \mathbf{f}(\mathbf{a}+t(\mathbf{b}-\mathbf{a}))$ es continua en $[0,1]$ y derivable en $(0,1)$ con derivada $\boldsymbol{\varphi}'(t) = d\mathbf{f}(\mathbf{a}+t(\mathbf{b}-\mathbf{a}))(\mathbf{b}-\mathbf{a})$. Para todo $t \in (0,1)$ se verifica

$$\|\boldsymbol{\varphi}'(t)\| \leq \|d\mathbf{f}(\mathbf{a}+t(\mathbf{b}-\mathbf{a}))\| \, \|\mathbf{b}-\mathbf{a}\| \leq M \, \|\mathbf{b}-\mathbf{a}\|$$

y según 4.8, $\|\boldsymbol{\varphi}(1) - \boldsymbol{\varphi}(0)\| \leq M \, \|\mathbf{b}-\mathbf{a}\|$ es decir $\|\mathbf{f}(\mathbf{b}) - \mathbf{f}(\mathbf{a})\| \leq M \, \|\mathbf{b}-\mathbf{a}\|$. ■

Un subconjunto A del espacio normado $(E, \| \ \|)$ se dice que es *convexo* cuando para cada par de puntos $\mathbf{x}, \mathbf{y} \in A$, el segmento $[\mathbf{x}, \mathbf{y}]$ está contenido en A.

Corolario 5.23 *Sea Ω un abierto convexo del espacio normado $(E, \| \ \|)$ y $\mathbf{f} : \Omega \to F$ una aplicación diferenciable tal que $\|d\mathbf{f}(\mathbf{x})\| \leq M$ para todo $\mathbf{x} \in \Omega$. Entonces $\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M \, \|\mathbf{y}-\mathbf{x}\|$ para cada $\mathbf{x}, \mathbf{y} \in \Omega$. En particular, si $d\mathbf{f}(\mathbf{x}) = 0$ para todo $\mathbf{x} \in \Omega$, entonces $\mathbf{f}$ es constante.*

Dem: Basta aplicar la proposición 5.22 a cada segmento $[\mathbf{x}, \mathbf{y}] \subset \Omega$. Si $d\mathbf{f}(\mathbf{x}) = 0$ para todo $\mathbf{x} \in \Omega$, podemos tomar $M = 0$ y se obtiene que $\mathbf{f}$ es constante. ■

Corolario 5.24 *Sea Ω un abierto conexo del espacio normado $(E, \| \ \|)$ y $\mathbf{f} : \Omega \to F$ una aplicación diferenciable con valores en el espacio normado $(F, \| \ \|)$. Si $d\mathbf{f}(\mathbf{x}) = 0$ para todo $\mathbf{x} \in \Omega$ entonces $\mathbf{f}$ es constante.*

Dem: Fijado $\mathbf{a} \in \Omega$, como $\mathbf{f}$ es continua $A = \{\mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a})\}$ es un subconjunto cerrado para la topología relativa de Ω. A no es vacío, pues $\mathbf{a} \in A$, y bastará ver que A es un subconjunto abierto del espacio conexo Ω para concluir que $A = \Omega$, y con ello que $\mathbf{f}$ es constante. Efectivamente, si $\mathbf{b} \in A$ existe $r > 0$ tal que $B(\mathbf{b}, r) \subset \Omega$. La bola $B(\mathbf{b}, r)$ es convexa y aplicando el corolario 5.23 se obtiene que $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{b})$ para todo $\mathbf{x} \in B(\mathbf{b}, r)$. Como $\mathbf{b} \in A$ se cumple que $\mathbf{f}(\mathbf{b}) = \mathbf{f}(\mathbf{a})$, luego $B(\mathbf{b}, r) \subset A$ y con esto queda probado que A es abierto. ■

5.4. Gradiente

Cada vector $\mathbf{z} = (z_1, z_2, \cdots, z_n) \in \mathbb{R}^n$ define una aplicación lineal $L_{\mathbf{z}} : \mathbb{R}^n \to \mathbb{R}$ mediante la fórmula $L_{\mathbf{z}}(\mathbf{h}) = \sum_{k=1}^n z_k h_k = \langle \ \mathbf{z} \mid \mathbf{h} \ \rangle$.

Recíprocamente, si $L : \mathbb{R}^n \to \mathbb{R}$ es lineal el vector $\mathbf{z} = (z_1, z_2, \cdots, z_n) \in \mathbb{R}^n$ definido por $z_k = L(\mathbf{e}_k)$, cumple $L = L_{\mathbf{z}}$ pues

$$L(\mathbf{h}) = L\left(\sum_{k=1}^n h_k \mathbf{e}_k\right) = \sum_{k=1}^n h_k L(\mathbf{e}_k) = \sum_{k=1}^n h_k z_k$$

Es decir, cada aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}$ se puede identificar con el único vector $\mathbf{z} \in \mathbb{R}^n$ que verifica $L(\mathbf{h}) = \langle \ \mathbf{z} \mid \mathbf{h} \ \rangle$ para todo $\mathbf{h} \in \mathbb{R}^n$. Cuando L es la diferencial de una función se define

Definición 5.25 *Sea $f : \Omega \to \mathbb{R}$ definida en un abierto $\Omega \subset \mathbb{R}^n$. Si f es diferenciable en $\mathbf{a} \in \Omega$ el gradiente de f en $\mathbf{a}$, denotado $\nabla f(\mathbf{a})$, es el vector de $\mathbb{R}^n$ asociado a su diferencial $df(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}$. Es decir, $\nabla f(\mathbf{a})$ es el único vector de $\mathbb{R}^n$ que verifica $df(\mathbf{a})\mathbf{h} = \langle\ \nabla f(\mathbf{a}) \mid \mathbf{h}\rangle$ para todo $\mathbf{h} \in \mathbb{R}^n$. Sus componentes vienen dadas por $df(\mathbf{a})\mathbf{e}_k = D_k f(\mathbf{a})$, $1 \leq k \leq n$, luego*

$$\nabla f(\mathbf{a}) = (D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \cdots, D_n f(\mathbf{a})) \in \mathbb{R}^n$$

El interés de la noción de gradiente lo muestran las dos proposiciones siguientes, que son la base de sus interpretaciones físicas y geométricas.

Proposición 5.26 *Sea $f : \Omega \to \mathbb{R}$ diferenciable en $\mathbf{a} \in \Omega$ con $\nabla f(\mathbf{a}) \neq (0, \cdots, 0)$. El máximo valor de la derivada direccional $D_{\mathbf{u}} f(\mathbf{a})$ con $\|\mathbf{u}\|_2 = 1$, se alcanza en la dirección del gradiente, $\mathbf{w} = \nabla f(\mathbf{a}) / \|\nabla f(\mathbf{a})\|_2$, y su valor es $\|\nabla f(\mathbf{a})\|_2$*

Dem: Es consecuencia inmediata de la definición de gradiente y de la desigualdad de Cauchy-Schwarz, pues si $\mathbf{u} \in \mathbb{R}^n$ y $\|\mathbf{u}\|_2 = 1$ se verifica

$$D_{\mathbf{u}} f(\mathbf{a}) = df(\mathbf{a})\mathbf{u} = \langle\ \nabla f(\mathbf{a}) \mid \mathbf{u}\ \rangle \leq |\langle\ \nabla f(\mathbf{a}) \mid \mathbf{u}\ \rangle| \leq \|\nabla f(\mathbf{a})\|_2 \|\mathbf{u}\|_2 = \|\nabla f(\mathbf{a})\|_2$$

Cuando $\mathbf{u} = \mathbf{w} = \nabla f(\mathbf{a}) / \|\nabla f(\mathbf{a})\|_2$ resulta

$$D_{\mathbf{w}} f(\mathbf{a}) = \langle\ \nabla f(\mathbf{a}) \mid \mathbf{w}\ \rangle = \frac{1}{\|\nabla f(\mathbf{a})\|_2} \langle\ \nabla f(\mathbf{a}) \mid \nabla f(\mathbf{a})\ \rangle = \|\nabla f(\mathbf{a})\|_2$$

■

Proposición 5.27 *Si $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, es diferenciable en $\mathbf{a} \in \Omega$ y $f(\mathbf{a}) = c$ entonces el vector gradiente $\nabla f(\mathbf{a})$ es ortogonal al conjunto de nivel $N_c = \{\mathbf{x} \in \Omega : f(\mathbf{x}) = c\}$ en el punto $\mathbf{a}$. (esto significa que es ortogonal a cada vector tangente a N_c en $\mathbf{a}$).*

Dem: Si $\mathbf{u} \in \mathbb{R}^n$ es tangente a N_c en el punto $\mathbf{a}$ existe un camino $\boldsymbol{\gamma} : (-r, r) \to N_c$, con $\boldsymbol{\gamma}(0) = \mathbf{a}$ y $\boldsymbol{\gamma}'(0) = \mathbf{u}$. La función $\varphi(t) = f(\boldsymbol{\gamma}(t))$ es constante (=c) en el intervalo $(-r, r)$ luego, en virtud de la regla de la cadena 5.21, se obtiene

$$0 = \varphi'(0) = df(\boldsymbol{\gamma}(0))\boldsymbol{\gamma}'(0) = df(\mathbf{a})\mathbf{u} = \langle\ \nabla f(\mathbf{a}) \mid \mathbf{u}\ \rangle$$

■

Interpretación física. Si pensamos que $T = f(x, y, z)$ es la temperatura del punto $(x, y, z) \in \Omega \subset \mathbb{R}^3$, según la proposición 5.26, la dirección según la cual varia más rápidamente la temperatura es la del gradiente $\nabla f(\mathbf{a})$, y la proposición 5.27, nos dice que esta dirección es ortogonal a la 'superficie' isoterma que pasa por $\mathbf{a}$.

Para una interpretación geométrica podemos pensar que la función de dos variables $z = f(x, y)$ proporciona la altura del terreno sobre el punto (x, y) de una región plana $\Omega \subset \mathbb{R}^2$. En este caso, si $\mathbf{a} \in \Omega$ es un punto donde h es diferenciable, la pendiente del terreno es máxima según la dirección del gradiente $\nabla f(\mathbf{a})$, que es

ortogonal a la 'curva' de nivel que pasa por $\mathbf{a}$.
NOTA: Utilizando el teorema B.8 se puede dar una definición general e intrínseca de gradiente (donde no intervienen las derivadas parciales) que sirve para el caso infinito dimensional: Sea $f : \Omega \to \mathbb{R}$ definida en un abierto Ω de un espacio normado completo $(E, \| \ \|)$ cuya norma procede de un producto escalar. Si f es diferenciable en $\mathbf{a}$, en virtud del teorema B.8 se define $\nabla f(\mathbf{a})$ como el único vector de E que cumple $\langle \nabla f(\mathbf{a}) \mid \mathbf{u} \rangle = df(\mathbf{a})\mathbf{u}$ para todo $\mathbf{u} \in E$. Es claro que, en este contexto más general, siguen valiendo las proposiciones 5.26 y 5.27.

5.5. Espacio tangente

En esta sección se establece que para una aplicación $\mathbf{f}$ diferenciable en un punto $\mathbf{a}$ hay una definición natural de espacio tangente a su gráfica en el punto $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$ ya que, en este caso, el conjunto de los vectores tangentes a la gráfica en ese punto es un espacio vectorial de dimensión igual al número de variables de la función. También se estudia la existencia de espacio tangente para conjuntos de nivel de aplicaciones diferenciables, y para conjuntos que se pueden parametrizar mediante una aplicación diferenciable.

Proposición 5.28 *Sea* $M = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{x} \in \Omega\}$ *la gráfica de* $\mathbf{f} : \Omega \to \mathbb{R}^{n-k}$*, definida en un abierto* $\Omega \subset \mathbb{R}^k$*. Si* $\mathbf{f}$ *es diferenciable en* $\mathbf{a}$ *y* $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$ *entonces*

$$T_{\mathbf{p}}(M) = \{(\mathbf{u}, d\mathbf{f}(\mathbf{a})\mathbf{u}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{u} \in \mathbb{R}^k\}$$

luego $T_{\mathbf{p}}(M)$ *es un subespacio vectorial k-dimensional de* $\mathbb{R}^n$*.*

DEM: Si $\mathbf{w} \in T_{\mathbf{p}}(M) \subset \mathbb{R}^n$, según la definición de vector tangente existe una función $\boldsymbol{\gamma} : (-r, r) \to \mathbb{R}^n$, derivable en 0, tal que $\gamma(-r, r) \subset M$, $\boldsymbol{\gamma}(0) = \mathbf{p}$ y $\boldsymbol{\gamma}'(0) = \mathbf{w}$.

En términos de las componentes de $\gamma(t) = (\boldsymbol{\gamma}_1(t), \boldsymbol{\gamma}_2(t))$, y $\mathbf{w} = (\mathbf{u}, \mathbf{v})$, en $\mathbb{R}^k \times \mathbb{R}^{n-k}$, las condiciones anteriores se escriben así:

$$\boldsymbol{\gamma}_2(t) = \mathbf{f}(\boldsymbol{\gamma}_1(t)) \text{ si } |t| < r, \quad \boldsymbol{\gamma}_1(0) = \mathbf{a}, \ \boldsymbol{\gamma}_2(0) = \mathbf{f}(\mathbf{a}), \quad \boldsymbol{\gamma}_1'(0) = \mathbf{u}, \ \boldsymbol{\gamma}_2'(0) = \mathbf{v}$$

En virtud de la regla de la cadena $\boldsymbol{\gamma}_2'(0) = d\mathbf{f}(\boldsymbol{\gamma}_1(0))\boldsymbol{\gamma}_1'(0)$, luego $\mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}$.

Recíprocamente, si $\mathbf{w} = (\mathbf{u}, \mathbf{v}) \in \mathbb{R}^k \times \mathbb{R}^{n-k}$ y $\mathbf{v} = d\mathbf{f}(\mathbf{a})\mathbf{u}$, podemos considerar la trayectoria $\boldsymbol{\gamma}_{\mathbf{u}}(t) = (\mathbf{a} + t\mathbf{u}, \mathbf{f}(\mathbf{a} + t\mathbf{u}))$, definida en un entorno de 0, y contenida en M. Como $\boldsymbol{\gamma}_{\mathbf{u}}(0) = \mathbf{p}$, $\boldsymbol{\gamma}_{\mathbf{u}}'(0) = (\mathbf{u}, d\mathbf{f}(\mathbf{a})\mathbf{u}) = (\mathbf{u}, \mathbf{v}) = \mathbf{w}$, resulta $\mathbf{w} \in T_{\mathbf{p}}(M)$. ■

En las condiciones de la proposición 5.28 se dice que $T_{\mathbf{p}}(M)$ es el *espacio vectorial tangente* a M en $\mathbf{p}$. Si los vectores de este espacio vectorial se colocan con origen en $\mathbf{p}$, sus extremos forman $\mathbf{p} + T_{\mathbf{p}}(M)$. Esto es el *espacio afín tangente a* M *en* $\mathbf{p}$.
Si un elemento genérico de $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$ lo representamos en la forma $(\mathbf{x}, \mathbf{y})$ con $\mathbf{x} = (x_1, x_2, \cdots x_k)$, $\mathbf{y} = (y_1, y_2, \cdots y_{n-k})$, el espacio afín tangente a la gráfica de $\mathbf{f}$ en $\mathbf{p} = (\mathbf{a}, \mathbf{b})$, $(\mathbf{b} = \mathbf{f}(\mathbf{a}))$ viene dado por

$$\{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^k \times \mathbb{R}^{n-k} : \mathbf{y} - \mathbf{b} = d\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})\}$$

Si $m = n - k$, sus ecuaciones son:

$$\begin{cases} y_1 - b_1 = & D_1 f_1(\mathbf{a})(x_1 - a_1) + D_2 f_1(\mathbf{a})(x_2 - a_2) + \cdots\cdots + D_k f_1(\mathbf{a})(x_k - a_k) \\ y_2 - b_2 = & D_1 f_2(\mathbf{a})(x_1 - a_1) + D_2 f_2(\mathbf{a})(x_2 - a_2) + \cdots\cdots + D_k f_2(\mathbf{a})(x_k - a_k) \\ \quad\cdots\cdots & \\ y_m - b_m = & D_1 f_m(\mathbf{a})(x_1 - a_1) + D_2 f_m(\mathbf{a})(x_2 - a_2) + \cdots\cdots + D_k f_m(a)(x_k - a_k) \end{cases}$$

En particular, cuando $n = 3$, $k = 2$, $p = 1$, y $\mathbf{a} = (x_0, y_0)$, los vectores tangentes a la gráfica $G(f)$ en $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$ forman un subespacio vectorial de $\mathbb{R}^3$ de dimensión 2, y en este caso el plano afín tangente a $G(f)$ en $\mathbf{p}$ es el de ecuación $z - z_0 = df(x_0, y_0)(x - x_0, y - y_0)\}$, es decir

$$z - z_0 = D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0)$$

de modo que $(D_1 f(x_0, y_0), D_2 f(x_0, y_0), -1)$ es un vector normal al plano tangente.

Espacio tangente a un conjunto de nivel. Ahora consideramos el problema de la existencia de plano tangente a un conjunto de nivel $M = \{\mathbf{x} \in \Omega : g(x, y, z) = c\}$ en un punto $\mathbf{p} \in M$, donde $g : \Omega \to \mathbb{R}$ es diferenciable. Según la siguiente proposición siempre se verifica $T_{\mathbf{p}}(M) \subset \{(u_1, u_2, u_3) : D_1 f(\mathbf{p}) u_1 + D_2 f(\mathbf{p}) u_2 + D_3 f(\mathbf{p}) u_3 = 0\}$.

Proposición 5.29 *Sea* $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\} \subset \mathbb{R}^n$ *un conjunto de nivel de* $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$*, definida en un abierto* $\Omega \subset \mathbb{R}^n$*. Si* $\mathbf{g}$ *es diferenciable en* $\mathbf{p} \in M$ *entonces* $T_{\mathbf{p}}(M) \subset \{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$.

Dem: Si $\mathbf{u} \in T_{\mathbf{p}}(M)$, según la definición de vector tangente existe $\boldsymbol{\gamma} : (-r, r) \to \mathbb{R}^n$ tal que

$$\boldsymbol{\gamma}(t) \in M \text{ si } |t| < r, \quad \boldsymbol{\gamma}(0) = \mathbf{p}, \quad \boldsymbol{\gamma}'(0) = \mathbf{u}$$

Como $\mathbf{g} \circ \boldsymbol{\gamma}$ es constante resulta $\mathbf{0} = (\mathbf{g} \circ \boldsymbol{\gamma})'(0) = d\mathbf{g}(\boldsymbol{\gamma}(0))\boldsymbol{\gamma}'(0) = d\mathbf{g}(\mathbf{p})\mathbf{u}$. ■

nota: En general no es posible garantizar que en la proposición 5.29 se cumpla la igualdad ni tampoco que $T_{\mathbf{p}}(M)$ sea espacio vectorial. Esto se pone de manifiesto con el siguiente ejemplo.

Ejemplo 5.30 *Sea* $M = \{(x, y, z) : x^2 + y^2 - z^2 = 0\}$ *un cono de revolución de eje* z*, con vértice en* $\mathbf{p} = (0, 0, 0)$*. Entonces* $T_{\mathbf{p}}(M) = \{(u_1, u_2, u_3) : u_3^2 = u_1^2 + u_2^2\}$.

Dem: Si $\mathbf{u} = (u_1, u_2, u_3) \in T_p(M)$, según la definición de vector tangente existe $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}^3$, $\gamma = (\gamma_1, \gamma_2, \gamma_3)$, derivable en $t = 0$ tal que $\gamma_3(t)^2 = \gamma_1(t)^2 + \gamma_2(t)^2$, $\gamma_j(0) = 0$, y $\gamma_j'(0) = u_j$ para $1 \leq j \leq 3$. En virtud de la definición de derivada

$$\gamma_j(t) = t[u_j + \epsilon_j(t)], \quad \text{donde} \quad \lim_{t \to 0} \epsilon_j(t) = 0$$

Se sigue que $(u_3 + \epsilon_3(t))^2 = (u_1 + \epsilon_1(t))^2 + (u_2 + \epsilon_2(t))^2$ y pasando al límite cuando $t \to 0$ se obtiene que $u_3^2 = u_1^2 + u_2^2$.

Recíprocamente, todo vector $\mathbf{u} = (u_1, u_2, u_3) \in \mathbb{R}^3$ que cumpla $u_3^2 = u_1^2 + u_2^2$ es tangente a M en $(0,0,0)$ pues la función $\gamma(t) = t\mathbf{u}$ cumple los requisitos que intervienen en la definición de vector tangente: $\gamma(t) \in M$, para todo $t \in (-1,1)$, $\gamma(0) = (0,0,0)$ y $\gamma'(0) = \mathbf{u}$. ■

En las condiciones de la proposición 5.29, si ocurre que $T_{\mathbf{p}}(M) \subset \mathbb{R}^n$ es un subespacio vectorial de dimensión k diremos que existe espacio tangente al conjunto de nivel M en el punto $\mathbf{p} \in M$. En este caso el espacio afín tangente a M en $\mathbf{p}$ es

$$\{\mathbf{x} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})(\mathbf{x} - \mathbf{p}) = \mathbf{0}\}$$

y si $m = n - k$, sus ecuaciones son

$$\begin{cases} D_1 g_1(\mathbf{p})(x_1 - p_1) + D_2 g_1(\mathbf{p})(x_2 - p_2) + \cdots + D_n g_1(\mathbf{p})(x_n - p_n) = 0 \\ D_1 g_2(\mathbf{p})(x_1 - p_1) + D_2 g_2(\mathbf{p})(x_2 - p_2) + \cdots + D_n g_2(\mathbf{p})(x_n - p_n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ D_1 g_m(\mathbf{p})(x_1 - p_1) + D_2 g_m(\mathbf{p})(x_2 - p_2) + \cdots + D_n g_m(\mathbf{p})(x_n - p_n) = 0 \end{cases}$$

Obsérvese que el espacio tangente a una gráfica $G(\mathbf{f})$ aparece como caso particular de éste considerando la función $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{x}) - \mathbf{y}$, con $\mathbf{x} \in \mathbb{R}^k$, $\mathbf{y} \in \mathbb{R}^{n-k}$.

Cuando sólo hay una ecuación $(k = n-1)$, se dice que $M = \{\mathbf{x} \in \Omega : g(\mathbf{x}) = c)\}$ es una 'hipersuperficie' de nivel. Ahora, el espacio tangente $T_{\mathbf{p}}(M)$ es el hiperplano ortogonal al vector $\nabla g(\mathbf{p}) = (D_1 g(\mathbf{p}), D_2 g(\mathbf{p}), \cdots D_n g(\mathbf{p})$. El hiperplano afín tangente a M en $\mathbf{p}$, es el de ecuación

$$D_1 g(\mathbf{p})(x_1 - p_1) + D_2 g(\mathbf{p})(x_2 - p_2) + \cdots + D_n g(\mathbf{p})(x_n - p_n) = 0$$

En el caso particular de $n = 3$, $M = \{(x,y,z) \in \Omega : g(x,y,z) = c\}$ es la superficie de nivel de una función $g : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^3$, y diferenciable en $\mathbf{p} \in M$. Siempre ocurre que el vector $\nabla g(\mathbf{p})$ es ortogonal al conjunto $T_{\mathbf{p}}(M)$. Cuando $T_{\mathbf{p}}(M) \subset \mathbb{R}^3$ sea un subespacio vectorial de dimensión 2 se dirá que existe el plano tangente a la 'superficie' de nivel M en el punto $\mathbf{p} \in M$. El correspondiente plano afín tangente es el de ecuación

$$D_1 g(\mathbf{p})(x - x_0) + D_2 g(\mathbf{p})(y - y_0) + D_3 g(\mathbf{p})(z - z_0) = 0$$

Obsérvese que el caso del plano tangente a una 'superficie' explícita $z = f(x,y)$ es un caso particular cuando $c = 0$, $g(x,y,z) = f(x,y) - z$ y $\mathbf{p} = (x_0, y_0, f(x_0, y_0))$.

Si $U_{\mathbf{p}}$ es un entorno de $\mathbf{p} \in M$, es claro que $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(M \cap U_{\mathbf{p}})$. Combinando esta observación con la proposición 5.28 se obtiene que existe plano tangente a una superficie de nivel $M = \{(x,y,z) \in \Omega : g(x,y,z) = c\}$ en un punto $\mathbf{p} \in M$ si se cumple la siguiente condición:

(T) Existe un entorno abierto $U_{\mathbf{p}}$ de $\mathbf{p}$ tal que $M \cap U_p$ se puede representar de alguna de las tres formas siguientes
a) $\{(x, y, f_3(x, y)) : (x, y) \in A_{12}\}$*;*
b) $\{(x, f_2(x, z), z) : (x, z) \in A_{13}\}$*;*
c) $\{(f_1(y, z), y, z) : (y, z) \in A_{23}\}$*.*
donde A_{ij} es un entorno abierto de (p_i, p_j) y $f_k : A_{ij} \to \mathbb{R}$ es diferenciable en (p_i, p_j) con $f_k(p_i, p_j) = p_k$. ((i, j, k) es una permutación de $(1, 2, 3)$).

En otras palabras, la condición (T) significa que después de realizar una permutación de las variables (x, y, z), la superficie de nivel M se puede representar localmente, en un entorno de $\mathbf{p}$, como la gráfica de una función $z = f(x, y)$ diferenciable en (p_1, p_2).

Si se cumple la condición (T), y alguna de las derivadas parciales $D_j g(\mathbf{p})$ no es nula, entonces los dos conjuntos que intervienen en la inclusión

$$T_{\mathbf{p}}(M) \subset \{(u_1, u_2, u_3) : D_1 g(\mathbf{p})u_1 + D_2 g(\mathbf{p})u_2 + D_3 g(\mathbf{p})u_3 = 0\}$$

son subespacios vectoriales de dimensión 2 y por lo tanto son iguales

Es claro que una esfera $S = \{(x, y, z) : x^2 + y^2 + z^2 = R^2\}$ cumple la condición (T) en cualquier punto $\mathbf{p} \in S$ y queda justificado así que el conjunto de los vectores tangentes a la esfera S en un punto $\mathbf{p} \in S$ es un espacio vectorial de dimensión 2.

Dejamos al cuidado del lector la formulación de la condición (T) para el caso de una función de n variables reales. Anticipamos que, en virtud del teorema de la función implícita 8.16 la condición (T) se cumple cuando $g : \Omega \to \mathbb{R}$, tiene derivadas parciales continuas en un entorno de $\mathbf{p}$ y no es nulo el gradiente $\nabla g(\mathbf{p})$.

La siguiente proposición extiende a una situación más general la condición suficiente para la existencia de espacio tangente a un conjunto de nivel. Para simplificar su enunciado hacemos el siguiente convenio: Si $I = \{i_1, i_2, \cdots, i_k\} \subset \{1, 2, \cdots, n\}$, para cada $\mathbf{x} \in \mathbb{R}^n$ denotaremos por $\mathbf{x}_I$ el vector de $\mathbb{R}^k = \mathbb{R}^I$, definido por $\mathbf{x}_I = (x_{i_1}, x_{i_2}, \cdots x_{i_k})$. Si $J = \{j_1, j_2, \cdots, j_{n-k}\} = \{1, 2, \cdots, n\} \setminus I$, podemos identificar $\mathbb{R}^n$ con $\mathbb{R}^I \times \mathbb{R}^J$ mediante la biyección natural $\mathbf{x} \leftrightarrow (\mathbf{x}_I, \mathbf{x}_J)$.

Proposición 5.31 *Sea $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\}$ donde $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$, está definida en un abierto $\Omega \subset \mathbb{R}^n$ y es diferenciable en $\mathbf{p} \in M$. Se supone que existe un entorno $\Omega_{\mathbf{p}}$ de $\mathbf{p}$ tal que $M \cap U_{\mathbf{p}}$ es la gráfica de una aplicación de k variables reales $\mathbf{x}_J = \mathbf{f}(\mathbf{x}_I)$ (definida en un abierto de $\mathbb{R}^I$, con valores en $\mathbb{R}^J$) diferenciable en $\mathbf{p}_J$. Entonces $T_{\mathbf{p}}(M)$ es un subespacio vectorial de dimensión k.*

DEM: Es fácil ver que $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$. En virtud de 5.28 la hipótesis garantiza que $T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$ es un subespacio vectorial de dimensión k. ■

NOTA. El teorema de la función implícita, que se verá más adelante, asegura que si $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$ es de clase C^1 en $\Omega \subset \mathbb{R}^n$ y $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{c}\}$ entonces, en cada $\mathbf{p} \in M$ donde la matriz jacobiana $\mathbf{g}'(\mathbf{p})$ tenga rango $n - k$ se cumple la condición requerida en la proposición anterior 5.31.

Espacio tangente a la imagen de una parametrización. Consideremos ahora un conjunto de la forma $M = \boldsymbol{\varphi}(\Omega)$, donde $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^n$, definida en un abierto $\Omega \subset \mathbb{R}^k$, es diferenciable en $\mathbf{a} \in \Omega$. Si $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$, sabemos que

$$d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k) = \{D_{\mathbf{u}}\boldsymbol{\varphi}(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^k\} \subset T_{\mathbf{p}}(M)$$

Si los vectores $D_1\boldsymbol{\varphi}(\mathbf{a}) ,\cdots, D_k\boldsymbol{\varphi}(\mathbf{a})$ son linealmente independientes, forman una base del subespacio $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k)$, que será de dimensión k. Cuando se verifique la igualdad $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k) = T_{\mathbf{p}}(M)$, diremos que existe espacio tangente a $M = \boldsymbol{\varphi}(\Omega)$ en el punto $\mathbf{p} \in M$. En este caso, el espacio afín tangente a M en $\mathbf{p}$, viene dado por

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{a} + d\boldsymbol{\varphi}(\mathbf{a})(t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \cdots + t_k\mathbf{e}_k) : (t_1, t_2, \cdots, t_k) \in \mathbb{R}^k\}$$

por lo que sus ecuaciones paramétricas son

$$\mathbf{x} = \mathbf{a} + t_1D_1\boldsymbol{\varphi}(\mathbf{a}) + t_2D_2\boldsymbol{\varphi}(\mathbf{a}) + \cdots + t_kD_k\boldsymbol{\varphi}(\mathbf{a}), \quad \text{con } (t_1, t_2, \cdots, t_k) \in \mathbb{R}^k$$

es decir,

$$\begin{cases} x_1 = a_1 + t_1D_1\varphi_1(\mathbf{a}) + \cdots + t_kD_k\varphi_1(\mathbf{a}) \\ x_2 = a_2 + t_1D_1\varphi_2(\mathbf{a}) + \cdots + t_kD_k\varphi_2(\mathbf{a}) \\ \cdots\cdots\cdots \\ x_n = a_n + t_1D_1\varphi_n(\mathbf{a}) + \cdots + t_kD_k\varphi_n(\mathbf{a}) \end{cases}$$

Obsérvese que el caso del espacio tangente a una gráfica $G(\mathbf{f})$ aparece como caso particular considerando la parametrización estandar

$$\boldsymbol{\varphi}(\mathbf{t}) = (\mathbf{t}, \mathbf{f}(\mathbf{t})) = (t_1, \cdots, t_k, f_1(t_1, \cdots, t_k), \cdots, f_{n-k}(t_1, \cdots, t_k))$$

En general, para un conjunto de la forma $M = \boldsymbol{\varphi}(\Omega)$, si $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ y $\boldsymbol{\varphi}$ es diferenciable en $\mathbf{a}$, sólo es posible asegurar la inclusión $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k) \subset T_{\mathbf{p}}(M)$ y puede ocurrir que $T_{\mathbf{p}}(M)$ no sea un espacio vectorial de dimensión n (véase el ejemplo 5.33). La siguiente proposición da una condición suficiente para la existencia de espacio tangente a un conjunto de esta forma.

Proposición 5.32 *Sea $M = \boldsymbol{\varphi}(\Omega)$ donde $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^n$ está definida en un abierto $\Omega \subset \mathbb{R}^k$, y es diferenciable en $\mathbf{a} \in \Omega$. Se supone que los vectores $D_1\boldsymbol{\varphi}(\mathbf{a}),\cdots,$ $D_k\boldsymbol{\varphi}(\mathbf{a})$ son linealmente independientes y que existe un entorno abierto $U_{\mathbf{p}} \subset \mathbb{R}^n$ de $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ tal que $M \cap U_{\mathbf{p}}$ se puede representar en forma implícita*

$$U_{\mathbf{p}} \cap M = \{\mathbf{x} \in U_{\mathbf{p}} : \mathbf{g}(\mathbf{x}) = 0\}$$

donde $\mathbf{g} : U_{\mathbf{p}} \to \mathbb{R}^{n-k}$ es diferenciable en $\mathbf{p}$ y el rango de la matriz jacobiana $\mathbf{g}'(\mathbf{p})$ es $n-k$. Entonces $T_{\mathbf{p}}(M)$ es un subespacio vectorial de dimensión k y por lo tanto existe el plano tangente a M en el punto $\mathbf{p}$.

DEM: Es fácil ver que $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(U_{\mathbf{p}} \cap M)$. Sabemos que se verifica

$$d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k) \subset T_{\mathbf{p}}(M) = T_{\mathbf{p}}(U_{\mathbf{p}} \cap M) \subset \{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$$

Como $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^k)$ y $\{\mathbf{u} \in \mathbb{R}^n : d\mathbf{g}(\mathbf{p})\mathbf{u} = \mathbf{0}\}$ son subespacios vectoriales de dimensión k, deben ser iguales, y por lo tanto coinciden con $T_{\mathbf{p}}(M)$. ■

NOTA En el capítulo 8 se verán condiciones suficientes para que se cumpla la condición requerida en la proposición 5.32 y en consecuencia, para que exista el plano tangente a $M = \boldsymbol{\varphi}(\Omega)$ en $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$.

Consideremos por último la situación más concreta de una 'superficie' paramétrica $M = \boldsymbol{\varphi}(\Omega)$, donde $\boldsymbol{\varphi} : \Omega \to \mathbb{R}^3$ está definida en un abierto $\Omega \subset \mathbb{R}^2$ y es diferenciable en $\mathbf{a} \in \Omega$. Si $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$, sabemos que

$$d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^2) = \{D_{\mathbf{u}}\boldsymbol{\varphi}(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^2\} \subset T_{\mathbf{p}}(M)$$

Si los vectores $D_1\boldsymbol{\varphi}(\mathbf{a})$, $D_2\boldsymbol{\varphi}(\mathbf{a})$ son linealmente independientes forman una base del subespacio $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^2)$. Cuando se cumpla la igualdad $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^2) = T_{\mathbf{p}}(M)$, y exista el plano tangente, en $\mathbf{p}$, a la 'superficie' paramétrica $M = \boldsymbol{\varphi}(\Omega)$, el plano afín tangente $\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = \mathbf{a} + d\boldsymbol{\varphi}(\mathbf{a})(s\mathbf{e}_1 + t\mathbf{e}_2) : (s,t) \in \mathbb{R}^2\}$ tendrá las ecuaciones paramétricas $\mathbf{x} = \mathbf{a} + sD_1\boldsymbol{\varphi}(\mathbf{a}) + tD_2\boldsymbol{\varphi}(\mathbf{a}), \quad$ con $(s,t) \in \mathbb{R}^2$, es decir

$$\begin{cases} x_1 = a_1 + sD_1\varphi_1(\mathbf{a}) + tD_2\varphi_1(\mathbf{a}) \\ x_2 = a_2 + sD_1\varphi_2(\mathbf{a}) + tD_2\varphi_2(\mathbf{a}) \\ x_3 = a_3 + sD_1\varphi_3(\mathbf{a}) + tD_2\varphi_3(\mathbf{a}) \end{cases}$$

El caso del plano tangente a una 'superficie' explícita $z = f(x,y)$ considerado anteriormente se obtiene como caso particular de este con la parametrización estandar $\boldsymbol{\varphi}(s,t) = (s,t,f(s,t))$. Obsérvese que las derivadas parciales de esta parametrización estandar proporcionan dos vectores tangentes linealmente independientes

$$D_1\boldsymbol{\varphi}(x_0,y_0) = (1,0,D_1f(x_0,y_0)), \quad D_2\boldsymbol{\varphi}(x_0,y_0) = (0,1,D_2f(x_0,y_0))$$

que, en este caso, forman una base del plano tangente $T_{\mathbf{p}}(M)$.

En general, para una superficie paramétrica $M = \boldsymbol{\varphi}(\Omega)$, si $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ y $\boldsymbol{\varphi}$ es diferenciable en $\mathbf{a}$, sólo es posible asegurar la inclusión $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^2) \subset T_{\mathbf{p}}(M)$ y puede ocurrir que $T_{\mathbf{p}}(M)$ no sea un plano, como se muestra en el siguiente ejemplo.

Ejemplo 5.33 *Sea* $\boldsymbol{\varphi} : \mathbb{R}^2 \to \mathbb{R}^3$ *definida por* $\boldsymbol{\varphi}(r,t) = (r\cos t, r\operatorname{sen} t, r)$ *y* $\mathbf{a} = (0,0)$. *En este caso* M *es un cono de revolución con vértice en* $(0,0,0)$. *Ya hemos visto en 5.30 que* $T_{\mathbf{p}}(M)$ *no es un espacio vectorial. Los tres vectores linealmente independientes* $(1,0,1)$, $(-1,0,1)$, $(0,1,1)$ *pertenecen a* $T_{\mathbf{p}}(M)$ *y sin embargo* $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R}^2)$ *es un subespacio de dimensión* 1 *porque* $d\boldsymbol{\varphi}(\mathbf{a})\mathbf{e}_1 = D_1\boldsymbol{\varphi}(\mathbf{a}) = (1,0,1)$ *y* $d\boldsymbol{\varphi}(\mathbf{a})\mathbf{e}_2 = D_2\boldsymbol{\varphi}(\mathbf{a}) = (0,0,0)$

5.6. Ejercicios resueltos

Ejercicio 5.34 *Demuestre que es uniformemente continua la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por*

$$f(x, y) = \frac{x^3 - y^3}{x^2 + y^2} \quad si \quad (x, y) \neq (0, 0); \; f(0, 0) = 0 \quad si \quad (x, y) = (0, 0)$$

solución

Lo demostraremos viendo que en todo punto $(x, y) \in \mathbb{R}^2$ existen las derivadas parciales $D_1 f(x, y), D_2 f(x, y)$, y las funciones $D_1 f$, $D_2 f$ están acotadas. Utilizando lo que se acaba de establecer se demuestra fácilmente que la función

$$g(x, y) = \frac{x^3}{x^2 + y^2} \quad \text{si} \quad (x, y) \neq (0, 0); \; g(0, 0) = 0 \quad \text{si} \quad (x, y) = (0, 0)$$

es uniformemente continua. Basta observar que en todo punto (x, y) existen las derivadas parciales

- $D_1 g(x, y) = \dfrac{x^4 + 3x^2 y^2}{(x^2 + y^2)^2}$ si $(x, y) \neq (0, 0); D_1 g(0, 0) = 1.$
- $D_2 g(x, y) = \dfrac{-2x^3 y}{(x^2 + y^2)^2}$ si $(x, y) \neq (0, 0); D_2 g(0, 0) = 0.$

Utilizando coordenadas polares $(x, y) = (r \cos\varphi, r \operatorname{sen}\varphi)$ se observa que para todo $(x, y) \in \mathbb{R}^2$ se cumple $|D_1 f(x, y)| \leq 4$ y $|D_2 f(x, y)| \leq 2$ luego g es uniformemente continua en $\mathbb{R}^2$. Se sigue que $f(x, y) = g(x, y) - g(y, x)$ también es uniformemente continua en $\mathbb{R}^2$. ■

Ejercicio 5.35 *Si* $f : \Omega \to \mathbb{R}$ *es la función definida en el ejemplo 5.7 compruebe que* $g(x, y) = x f(x, y)$ *no es uniformemente continua en* Ω*, pero tiene derivadas parciales acotadas.*

solución

$$D_1 g(x, y) = f(x, y) - x D_1 f(x, y); \quad D_2 g(x, y) = x D_2 f(x, y)$$

y con los valores de $D_1 f(x, y)$, $D_2 f(x, y)$ vistos en 5.7 se obtiene

$$D_1 g(x, y) = f(x, y) - \frac{xy}{x^2 + y^2}; \quad D_2 g(x, y) = \frac{x^2}{x^2 + y^2}$$

luego $|D_1 g(x, y)| \leq \pi + 1$, y $|D_2 g(x, y)| \leq 1$ para todo $(x, y) \in \Omega$.

La función g no es uniformemente continua en Ω, porque no admite una extensión continua a $\overline{\Omega} = \mathbb{R}^2$, ya que para $a < 0$ se cumple

$$\lim_{y \to 0+} g(a, y) = a\pi; \quad \lim_{y \to 0-} g(a, y) = -a\pi$$

■

Ejercicio 5.36 *Se considera la función de dos variables*

$$f(x,y) = yx^2 \operatorname{sen}(1/x) si\ x \neq 0; \quad f(0,y) = 0$$

a) Demuestre que f es diferenciable en todo punto $(x,y) \in \mathbb{R}^2$

b) Determine un punto $\mathbf{p}$ de la elipse $\{(x,y) : x^2 + \pi^2 y^2 = 1\}$ tal que un vector unitario tangente en $\mathbf{p}$ a la elipse proporcione el máximo valor de la derivada direccional $D_{\mathbf{u}} f(2/\pi, \pi)$.

SOLUCIÓN

a) En todo punto $(x,y) \in \mathbb{R}^2$ existe la derivada parcial $D_1 f(x,y)$:
$D_1 f(x,y) = 2xy \operatorname{sen}(1/x) - y\cos(1/y)$ si $x \neq 0$.
$D_1 f(0,y) = \lim_{h \to 0}[f(h,y) - f(0,y)]/h = 0$.

En todo punto $(x,y) \in \mathbb{R}^2$ existe la derivada parcial $D_2 f(x,y)$:
$D_2 f(x,y) = x^2 \operatorname{sen}(1/x)$ si $x \neq 0$.
$D_2 f(0,y) = \lim_{k \to 0}[f(0,y+k) - f(0,y)]/h = 0$.

Se comprueba que $D_2 f$ es continua en todo punto $(x,y) \in \mathbb{R}^2$. Entonces, según la condición suficiente de diferenciabilidad, f es diferenciable en todo punto. (Observación: Para aplicar la condición suficiente de diferenciabilidad basta que sea continua una de las dos derivadas parciales. En este caso se puede ver que $D_1 f(x,y)$ no es continua en los puntos de la forma $(0,y)$ con $y \neq 0$).
b) El valor máximo de la derivada direccional se alcanza en la dirección del vector $\nabla f(2/\pi, \pi) = (4, 4/\pi^2)$. Sea $g(x,y) = x^2 + \pi^2 y^2 - 1$. Si (a,b) es un punto de la elipse, los vectores tangentes a la elipse en (a,b) son los ortogonales a $\nabla g(a,b) = (2a, 2\pi^2 b)$.

Buscamos un punto (a,b) de la elipse tal que $(4, 4/\pi^2)$ sea ortogonal a $(2a, 2\pi^2 b)$, es decir, un punto que verifique $0 = a+b$; $a^2 + \pi^2 b^2 = 1$;. Como soluciones se obtienen los dos puntos de la elipse $(a, -a)$ y $(-a, a)$ con $a = \sqrt{1/(1+\pi^2)}$.

Hay dos vectores unitarios $\mathbf{u}, -\mathbf{u}$ tangentes a la elipse en $(a,-a)$ (resp. $(-a,a)$) y proporciona la derivada direccional máxima el que tiene la misma dirección que $\nabla f(2/\pi, \pi)$. ■

Ejercicio 5.37 *Si $\mathbf{a} \in \mathbb{R}^n$, demuestre que la función $f : \mathbb{R}^n \to \mathbb{R}$, definida por*

$$f(\mathbf{x}) = \langle \mathbf{x} \mid \mathbf{a} \rangle e^{-\|\mathbf{x}\|_2}$$

alcanza en $\mathbb{R}^n$ un máximo y un mínimo absoluto. Calcule sus valores.

SOLUCIÓN

Como el caso $\mathbf{a} = \mathbf{0}$ es trivial suponemos en lo que sigue que $\mathbf{a} \neq \mathbf{0}$. En virtud de la desigualdad de Cauchy-Schwartz 2.2, para todo $\mathbf{x} \in \mathbb{R}^n$ se cumple

$$|f(\mathbf{x})| \leq \|\mathbf{a}\|_2 \|\mathbf{x}\|_2 e^{-\|\mathbf{x}\|_2}, \quad \text{luego} \quad \lim_{\|\mathbf{x}\|_2 \to +\infty} f(\mathbf{x}) = 0$$

Como $f(\mathbf{a}) > 0$, existe $R > 0$ tal que $\|\mathbf{x}\|_2 > R \Rightarrow |f(\mathbf{x})| < f(\mathbf{a}) = \|\mathbf{a}\|_2 e^{-\|\mathbf{a}\|_2}$.
Se sigue que $\|\mathbf{a}\|_2 \leq R$ y que el máximo absoluto que alcanza $|f|$ en un punto del compacto $K = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq R\}$ es realmente el máximo absoluto en $\mathbb{R}^n$.

Sea $\mathbf{b} \in K$ tal que $|f(\mathbf{b})| \geq |f(\mathbf{x})|$ para todo $\mathbf{x} \in \mathbb{R}^n$. Como $f(-\mathbf{b}) = -f(\mathbf{b})$, podemos suponer que $|f(\mathbf{b})| = f(\mathbf{b})$, luego para todo $\mathbf{x} \in \mathbb{R}^n$ se cumple $f(\mathbf{b}) \geq f(\mathbf{x})$. Es inmediato que si f alcanza un máximo absoluto en $\mathbf{b}$ entonces alcanza también un mínimo absoluto en $-\mathbf{b}$. Según la proposición 5.9, $\mathbf{b}$ es punto estacionario de f y debe ser solución del sistema

1) $\quad 0 = D_1 f(\mathbf{x}) \equiv \left(a_1 - \dfrac{x_1}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2}$

2) $\quad 0 = D_2 f(\mathbf{x}) \equiv \left(a_2 - \dfrac{x_2}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2}$

......

n) $\quad 0 = D_n f(\mathbf{x}) \equiv \left(a_n - \dfrac{x_n}{\|\mathbf{x}\|_2} \langle \mathbf{a} \mid \mathbf{x} \rangle \right) e^{-\|\mathbf{x}\|_2}$

Multiplicando la ecuación de la linea j por x_j y sumando se obtiene que los puntos estacionarios de f deben cumplir

$$0 = \langle \mathbf{a} \mid \mathbf{x} \rangle - \|\mathbf{x}\|_2 \langle \mathbf{a} \mid \mathbf{x} \rangle \qquad (*)$$

En particular, como el punto $\mathbf{b}$ donde f alcanza el máximo satisface (*) y cumple $\langle \mathbf{a} \mid \mathbf{b} \rangle = f(\mathbf{b}) e^{\|\mathbf{b}\|_2} > 0$, se obtiene que $\|\mathbf{b}\|_2 = 1$. Como $\mathbf{b}$ satisface el sistema de ecuaciones de los puntos estacionarios,

$$a_j = b_j \langle \mathbf{a} \mid \mathbf{b} \rangle, \quad 1 \leq j \leq n$$

Elevando al cuadrado y sumando para $1 \leq j \leq n$ se llega a la igualdad

$$\|\mathbf{a}\|_2^2 = \|\mathbf{b}\|_2^2 \langle \mathbf{a} \mid \mathbf{b} \rangle^2$$

Como $\|\mathbf{b}\|_2 = 1$ y $\langle \mathbf{a} \mid \mathbf{b} \rangle > 0$ resulta $\|\mathbf{a}\|_2 = \langle \mathbf{a} \mid \mathbf{b} \rangle$. Si en las ecuaciones de los puntos estacionarios sustituimos $\mathbf{x}$ por $\mathbf{b}$ y luego $\langle \mathbf{a} \mid \mathbf{b} \rangle$ por $\|\mathbf{a}\|_2$ se obtiene

$$0 = a_j - b_j \langle \mathbf{a} \mid \mathbf{b} \rangle = a_j - b_j \|\mathbf{a}\|_2$$

es decir, $\mathbf{b} = \mathbf{a}/\|\mathbf{a}\|_2$, luego el máximo absoluto es $f(\mathbf{a}/\|\mathbf{a}\|_2) = \|\mathbf{a}\|_2 / e$, y el mínimo absoluto $f(-\mathbf{a}/\|\mathbf{a}\|_2) = -\|\mathbf{a}\|_2 / e$. ■

Ejercicio 5.38 *Dados* $\mathbf{p}^1, \mathbf{p}^2, \cdots \mathbf{p}^m \in \mathbb{R}^n$ *demuestre que existe* $\mathbf{p} \in \mathbb{R}^n$ *donde la suma* $f(\mathbf{x}) = \sum_{k=1}^m \left\|\mathbf{x} - \mathbf{p}^k\right\|_2^2$ *alcanza un mínimo absoluto y calcule* $\mathbf{p}$.

solución

La desigualdad $f(\mathbf{x}) \geq \sum_{k=1}^m (\|\mathbf{x}\|_2 - \left\|\mathbf{p}^k\right\|_2)^2$, implica que $f(\mathbf{x})$ tiende hacia $+\infty$ cuando $\|\mathbf{x}\|_2 \to +\infty$, luego existe $R > 0$ tal que $\|\mathbf{x}\|_2 > R \Rightarrow f(\mathbf{x}) > f(0)$.

En algún punto $\mathbf{p}$ del compacto $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq R\}$ la función continua $f|_K$ alcanza un mínimo absoluto. Si $\|\mathbf{x}\| > R$ se cumple $f(\mathbf{x}) > f(0) \geq f(\mathbf{p})$, luego $\mathbf{f}(\mathbf{p})$ es realmente el mínimo absoluto de f en todo $\mathbb{R}^n$. según la proposición 5.9, $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ es un punto estacionario de f, es decir, satisface las ecuaciones

$$D_j f(\mathbf{x}) = 0, \quad 1 \le j \le n \quad (*)$$

Para calcular las derivadas parciales $D_j f$ escribimos $f(\mathbf{x}) = \sum_{k=1}^m f_k(\mathbf{x})$, donde $f_k(x_1, x_2, \cdots, x_n) = \sum_{j=1}^n (x_j - p_j^k)^2$. Resulta

$$D_j f(\mathbf{x}) = \sum_{k=1}^m D_j f_k(\mathbf{x}) = 2\sum_{k=1}^m (x_j - p_j^k) = 2(m x_j - \sum_{k=1}^m p_j^k); \quad 1 \le j \le n$$

Se sigue que $\mathbf{p}$ es la única solución de (*), luego $p_j = \frac{1}{m}\sum_{k=1}^m p_j^k$. Queda establecido así que f alcanza su mínimo absoluto en el punto $\mathbf{p} = \frac{1}{m}\sum_{k=1}^m \mathbf{p}^k$. ■

Ejercicio 5.39 *Demuestre que entre todos las cajas (con tapa) cuyo volumen es* 1 *litro hay una cuya superficie es mínima. Calcule sus dimensiones.*

solución

Si $x > 0$, $y > 0$, $z > 0$ son las dimensiones de la caja, expresadas en cm., hay que demostrar que área de la caja $S(x, y, z) = 2xy + 2xz + 2yz$ alcanza un mínimo absoluto cuando x, y, z cumplen la condición de ligadura $xyz = 1000$ cm^3.

Basta demostrar que la función de dos variables

$$f(x, y) = S(x, y, 1000/(xy) = 2xy + 2000/x + 2000/y$$

alcanza un mínimo absoluto en el abierto $U = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$. Haremos la demostración buscando un compacto $K \subset U$ donde podamos asegurar que el mínimo absoluto de $f|_K$ es realmente el mínimo absoluto de f en U.

Vemos a continuación que este requisito lo cumple

$$K = \{(x, y) \in \mathbb{R}^2 : x \ge 1, \ y \ge 1, \ xy \le 1000\} \subset U$$

que es compacto por ser cerrado y acotado (obsérvese que para $(x, y) \in K$ se cumple $0 \le y \le 1000$, $0 \le x \le 1000$). Cuando $(x, y) \in U \setminus K$ se verifica al menos una de las desigualdades $x < 1$, $y < 1$, $xy > 1000$, de donde se sigue que

$$f(x, y) = 2xy + 2000/x + 2000/y > 2000$$

Es fácil ver que existe $\mathbf{q} \in K$ tal que $f(\mathbf{q}) < 2000$ (por ejemplo $\mathbf{q} = (5, 10)$), luego el mínimo absoluto de la función continua f en el compacto K también es el mínimo absoluto de f en U. El punto $\mathbf{p} \in K$ donde f alcanza el mínimo absoluto debe satisfacer las ecuaciones

$$0 = D_1 f(x, y) = 2y - 2000/x^2; \qquad 0 = D_2 f(x, y) = 2x - 2000/y^2$$

que sólo tienen una sola solución, $x = y = 10$, luego $\mathbf{p} = (10, 10)$, lo que significa que la caja de superficie mínima es el cubo de lado 10 cm. ■

Ejercicio 5.40 *Calcule la mínima distancia entre la superficie $z = x^2 + y^2$ y el plano $x + 2y - z = 4$, justificando previamente que la mínima distancia se alcanza.*

DEM: Es fácil ver que el paraboloide de revolución $S = \{(x, y, z) : z = x^2 + y^2\}$ y el plano $P = \{(x, y, z) : x+2y-z = 4\}$ son disjuntos: Si hubiese algún punto (x, y), en su intersección, cumpliría $4 = x+2y-(x^2+y^2) = 5/4-(x-1/2)^2-(y-1)^2 \leq 5/4$!.

Como los conjuntos S y P son cerrados, pero ninguno es acotado no disponemos de un resultado topológico de carácter general que nos garantice que la distancia

$$d(S, P) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in S, \mathbf{y} \in P\}$$

se alcanza en una pareja de puntos $\mathbf{p} \in P$, $\mathbf{q} \in S$.

Sin embargo, en el caso que nos ocupa, lo podremos asegurar acudiendo a un resultado bien conocido de geometría elemental: La distancia de un punto (x, y, z) al plano $x + 2y - z = 4$ viene dada por la fórmula $D(x, y, z) = |x + 2y - z - 4|/\sqrt{6}$.

El problema lo podremos resolver demostrando primero que sobre el paraboloide S la función $D(x, y, z)$ alcanza un mínimo absoluto en algún punto $\mathbf{q} = (x_0, y_0, z_0)$ y calculando ese punto. Como la distancia de $\mathbf{q}$ al plano P se alcanza en algún $\mathbf{p} \in P$ (porque P es cerrado), quedará justificado que la distancia $d(S, P) = \|\mathbf{q} - \mathbf{p}\|_2$ se alcanza en los puntos $\mathbf{p} \in P, \mathbf{q} \in S$.

La función $D(x, y, z)$ alcanza un mínimo absoluto sobre el paraboloide $z = x^2+y^2$ si y sólo si la función de dos variables reales $f(x, y) = |x + 2y - (x^2 + y^2) - 4|$ alcanza un mínimo absoluto en algún punto. Completando cuadrados se observa que $x^2 + y^2 - x - 2y + 4 = (x - 1/2)^2 + (y - 1)^2 + 11/4 > 0$ para todo $(x, y) \in \mathbb{R}^2$, luego

$$f(x, y) = x^2 + y^2 - x - 2y + 4 = (x^2 + y^2)\left(\frac{4 - x - 2y}{x^2 + y^2} + 1\right)$$

es una función continua en $\mathbb{R}^2$ que cumple $\lim_{\|(x,y)\|_2 \to +\infty} f(x, y) = +\infty$, de donde se sigue, por un razonamiento estandar, que f alcanza un mínimo absoluto en algún (x_0, y_0), que debe satisfacer las ecuaciones $0 = D_1f(x, y) = D_2f(x, y)$. Como estas ecuaciones sólo tienen la solución $(1/2, 1)$, se sigue que $\mathbf{q} = (x_0, y_0, z_0) = (1/2, 1, 5/4)$ es el punto del paraboloide S que está más cerca del plano P. Finalmente, el cálculo del punto $\mathbf{p} \in P$ que está más cerca de $\mathbf{q}$ es un asunto elemental que dejamos al cuidado del lector (se puede hacer con los procedimientos usuales de geometría analítica, o con los métodos que estamos exponiendo). ■

Ejercicio 5.41 *En cada uno de los siguientes casos calcule los extremos absolutos de f sobre el compacto que se indica:*

a) $f(x, y) = x - x^2 - y^2$ $\quad K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$;

b) $f(x, y) = \operatorname{sen}(xy)$ $\quad K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$;

c) $f(x, y) = x^2 + 5y^3 - y$ $\quad K = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, y \geq 0\}$;

SOLUCIÓN

a) Un punto $\mathbf{p} \in K$ donde $f|_K$ alcanza un extremo absoluto puede estar en el interior de K o en su frontera, $C = \{(x, y) : x^2 + y^2 = 1\}$.

- Si ocurre lo primero, $\mathbf{p}$ debe satisfacer las ecuaciones $0 = D_1f(x, y) = 1 - 2x$;

$0 = D_2 f(x, y) = -2y$, que sólo tienen una solución $\mathbf{a} = (1/2, 0) \in K$.
- Si ocurre lo segundo $f|_C$ alcanza un extremo absoluto en $\mathbf{p}$. Los extremos absolutos de $f|_C$ son inmediatos en el caso que nos ocupa: Cuando $(x, y) \in C$ se cumple $f(x, y) = x - 1$ y es obvio que el mínimo (resp. el máximo) de $x - 1$ sobre la circunferencia C se alcanza en $\mathbf{b} = (-1, 0)$, (resp. $\mathbf{c} = (1, 0)$).

En definitiva, $f|_K$ alcanza sus extremos absolutos en algunos de los puntos $\mathbf{a}, \mathbf{b}, \mathbf{c}$. Como $f(\mathbf{a}) = 1/4$, $f(\mathbf{b}) = -2$, y $f(\mathbf{c}) = 0$, concluimos que el máximo absoluto es $1/4 = f(\mathbf{a})$, y el mínimo absoluto es $-2 = f(\mathbf{b})$.

b) El problema de calcular los extremos absolutos de $f(x, y) = \operatorname{sen}(xy)$ sobre K se simplifica calculando previamente los extremos absolutos de $g(x, y) = xy$ sobre K

$$\alpha = g(\mathbf{p}) = \min\{g(x, y) : (x, y) \in K\}; \quad \beta = g(\mathbf{q}) = \min\{g(x, y) : (x, y) \in K\}$$

que se alcanzarán en sendos puntos $\mathbf{p}, \mathbf{q} \in K$. En efecto, como $g(K)$ es conexo y compacto (por serlo K) también lo será $g(K)$, luego $g(K) = [\alpha, \beta]$. Entonces los extremos absolutos de $f(x, y) = \operatorname{sen} g(x, y)$ sobre K serán los extremos absolutos de la función $\operatorname{sen} t$ en el intervalo $[\alpha, \beta]$.

Los puntos de K donde $g|_K$ alcanza los extremos absolutos están en el interior de K, o en su frontera $C = \{(x, y) : x^2 + y^2 = 1\}$.

- Los que estén en el interior deben satisfacer las ecuaciones

$$0 = D_1 g(x, y) = y; \quad 0 = D_2 g(x, y) = x$$

Como su solución trivial $\mathbf{a} = (0, 0)$ es un punto interior a K, este será el primer candidato a punto de extremo absoluto (para la función $g|_K$).

- Los que estén en C serán puntos donde $g|_C$ alcance extremos absolutos. En el caso que nos ocupa estos extremos se pueden calcular fácilmente usando la parametrización de C, $x = \cos t$, $y = \operatorname{sen} t$. Con los recursos habituales del cálculo diferencial de funciones de una variable se obtiene que los extremos absolutos de $\varphi(t) = g(\cos t, \operatorname{sen} t) = \operatorname{sen} t \cos t$ sobre $[0, 2\pi]$ son

$$\alpha = \varphi(3\pi/4) = \varphi(7\pi/4) = -1/2; \qquad \beta = \varphi(\pi/4) = \varphi(5\pi/4) = 1/2$$

Se sigue de esto que el máximo absoluto de $g|_C$ es

$$\frac{1}{2} = g(\mathbf{b}) = g(-\mathbf{b}) \text{ donde } \mathbf{b} = (\sqrt{2}/2, \sqrt{2}/2)$$

y que el mínimo absoluto de de $g|_C$ es

$$-\frac{1}{2} = g(\mathbf{c}) = g(-\mathbf{c}), \text{ donde } \mathbf{c} = (+\sqrt{2}/2, -\sqrt{2}/2)$$

Como $g|_K$ alcanza sus extremos absolutos en puntos de la terna $\mathbf{a}, \mathbf{b}, \mathbf{c}$, concluimos que su mínimo absoluto es $-1/2 = g(\mathbf{c})$, y su máximo absoluto es $1/2 = \mathbf{g}(\mathbf{b})$.

Según lo indicado al principio, el mínimo y el máximo absoluto de $f|_K$ son, respectivamente, $-\operatorname{sen}(1/2) = f(\mathbf{c})$, $\operatorname{sen}(1/2) = f(\mathbf{b})$, pues estos son los valores extremos de $\operatorname{sen} t$ en el intervalo $[-1/2, 1/2]$.

c) Los puntos de K donde $f|_K$ alcanza los extremos absolutos estarán en el interior de K o en su frontera, $\partial K = S \cup L$ donde

$$L = \{(x, 0) : -1 \leq x \leq 1\}; \quad S = \{(x, y) : y \geq 0, x^2 + y^2 = 1\}$$

- De estos puntos, los que estén en el interior de K, deben satisfacer las ecuaciones

$$0 = D_1 f(x, y) = 2x; \quad 0 = D_2 f(x, y) = 15y^2 - 1$$

que sólo tienen una solución $\mathbf{a} = (0, 1/\sqrt{15})$ en el interior de K.
- Los otros puntos donde f puede alcanzar extremos absolutos son los puntos donde $f|_L$ y $f|_S$ alcanzan sus extremos absolutos.

Los extremos absolutos de $f|_L$, coinciden con los de $f(x, 0) = x^2$ sobre $[-1, 1]$, que son $0 = f(0, 0)$, y $1 = f(1, 0) = f(-1, 0)$.

Los extremos absolutos de $f|_S$, coinciden con los extremos absolutos del polinomio $p(y) = f(1 - y^2, y) = 5y^3 - y^2 - y + 1$ sobre el intervalo $[0, 1]$. Con las técnicas habituales para funciones de una variable se obtienen los valores extremos de p en $[0, 1]$, que son $p(1/3) = 20/27$ (mínimo), y $p(1) = 4$ (máximo). En S hay dos puntos con $y = 1/3$, que son $(\pm 2\sqrt{2}/3, 1/3)$, y un punto con $y = 1$, que es $(0, 1)$.

En definitiva, los extremos absolutos de $f|_K$ se alcanzan en algunos de los puntos de la siguiente lista de candidatos:

$$(0, 1/\sqrt{15}),\ (-1, 0),\ (0, 0),\ (1, 0),\ (2\sqrt{2}/3, 1/3),\ (-2\sqrt{2}/3, 1/3),\ (0, 1)$$

y evaluando f en estos puntos, se obtienen los extremos absolutos de $f|_K$:

$$-2/(3\sqrt{15}) = f(0, 1/\sqrt{15}); \qquad 4 = f(0, 1).$$

■

Ejercicio 5.42 *Sean $(E, \| \|), (F, \| \|)$ espacios normados. Una aplicación $\mathbf{f} : \Omega \to F$, definida en un abierto $\Omega \subset E$, es diferenciable en $\mathbf{a} \in \Omega$ si y sólo sí se cumplen las dos condiciones siguientes:*

a) Para cada $\mathbf{v} \in E$ existe $D_{\mathbf{v}} f(\mathbf{a})$ y la aplicación $\mathbf{v} \to D_{\mathbf{v}} \mathbf{f}(\mathbf{a})$ es lineal continua.

b) $\lim_{t \to 0} \dfrac{\mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a})}{t} = D_{\mathbf{v}} \mathbf{f}(\mathbf{a})$ *uniformemente en* $\{\mathbf{v} \in E : \|\mathbf{v}\| = 1\}$.

(es decir, la condición b) es lo que le falta a una aplicación que cumple a) para ser diferenciable).

solución

Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$, según 5.14, se cumple a). Sea $L = d\mathbf{f}(\mathbf{a})$ y $\boldsymbol{\alpha}(\mathbf{h}) = \mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - L(\mathbf{h})$ el error que se comete al aproximar el incremento $\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a})$

mediante el incremento de la diferencial $L(\mathbf{h}) = L(\mathbf{a}+\mathbf{h}) - L(\mathbf{a})$. Según la definición de diferencial $\lim_{\mathbf{h} \to 0} \boldsymbol{\alpha}(\mathbf{h})/\|\mathbf{h}\| = 0$, es decir, para cada $\epsilon > 0$ existe $\delta > 0$ tal que $\|\mathbf{h}\| < \delta \ \Rightarrow \|\boldsymbol{\alpha}(\mathbf{h})\| < \epsilon \|\mathbf{h}\|$. Si $|t| < \delta$, para cada $\mathbf{v} \in E$, con $\|\mathbf{v}\| = 1$, se cumple $\|\boldsymbol{\alpha}(t\mathbf{v})\| < \epsilon|t|$, luego $\lim_{t \to 0} \boldsymbol{\alpha}(t\mathbf{v})/t = 0$ uniformemente en $\{\mathbf{v} \in E : \|\mathbf{v}\| = 1\}$. Como $\boldsymbol{\alpha}(\mathbf{tv}) = \mathbf{f}(\mathbf{a} + t\mathbf{v}) - \mathbf{f}(\mathbf{a}) - tD_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ se obtiene b).

El recíproco es inmediato: En virtud de b) la aplicación lineal continua $L : E \to F$, $L(\mathbf{v}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ proporcionada por a) satisface la definición de diferencial. ■

5.7. Ejercicios propuestos

♢ **5.7.1** *Calcule, en un punto genérico, las derivadas parciales de las funciones*

a) x^{y^z}, *definida en* $\{(x,y,z) : x > 0, y > 0\}$.

b) $\operatorname{sen}(x^y + y^z + z^x)$, *definida en* $\{(x,y,z) : x > 0, y > 0, z > 0\}$.

c) $\operatorname{sen}(x \operatorname{sen}(y \operatorname{sen} z))$, *definida en todo* $\mathbb{R}^3$.

d) $\int_0^{x^2y^3} g(t)dt + \int_{x^2}^{x^2+z^2} g(t)dt$, *donde* $g : \mathbb{R} \to \mathbb{R}$ *es continua.*

♢ **5.7.2** *Demuestre que la función*

$$f(x,y) = \frac{x^3 + 2xy^2}{x^2 + y^2} \quad si\ (x,y) \neq (0,0), \quad f(0,0) = 0$$

es uniformemente continua en $\mathbb{R}^2$.

♢ **5.7.3** *Demuestre que la función* $f(x,y) = 1/(1 + x^2 + y^2)$ *es uniformemente continua en todo* $\mathbb{R}^2$. *Estudie si sus derivadas parciales alcanzan extremos absolutos en todo el plano y obtenga la mejor cota de* $|D_1 f(x.y)|$, $|D_2 f(x,y)|$.

♢ **5.7.4** *Demuestre que la función*

$$f(x,y) = \frac{xy}{(x+y)(1+x)(1+y)}$$

es uniformemente continua en el abierto $\{(x,y) : x > 0, y > 0\}$ *y alcanza un máximo absoluto en este abierto. Obténgalo.*

♢ **5.7.5** *Determine los valores de* $n \in \mathbb{N}$ *para los que es uniformemente continua la función* $f : \mathbb{R}^2 \to \mathbb{R}$, *definida por*

$$f(x,y) = \frac{y^n}{x^2 + y^2} \quad si\ (x,y) \neq (0,0); \quad f(0,0) = 0 \quad (n \in \mathbb{N})$$

♢ **5.7.6** *Sea* $f : \Omega \to \mathbb{R}$ *definida en el abierto* $\Omega = \mathbb{R}^2 \setminus \{(x,0) : x \geq 0\}$ *tal que en cada* $(x,y) \in \Omega$ *existe y es nula la derivada parcial* $D_1 f(x,y) = 0$. *Demuestre que* f *no depende de la primera variable. ¿Se obtiene un resultado análogo para funciones* $f : \Omega \to \mathbb{R}$ *tales que en cada* $(x,y) \in \Omega$ *existe y es nula la derivada parcial* $D_2 f(x,y) = 0$?.

♢ **5.7.7** *Obtenga los vectores* $\mathbf{v} \in \mathbb{R}^3$ *para los que existe la derivada* $D_{\mathbf{v}} f(1,-1,0)$ *de la función* $f(x,y,z) = |x + y + z|$.

♢ **5.7.8** *Sea* $A \subset \mathbb{R}^n$ *un abierto acotado y* $f : \overline{A} \to \mathbb{R}$ *una función continua, diferenciable en cada* $\mathbf{x} \in A$, *tal que* $f(\mathbf{x}) = 0$ *para todo* $\mathbf{x} \in \partial A$. *Demuestre que existe* $\mathbf{a} \in A$ *tal que* $\nabla f(\mathbf{a}) = 0$.

♢ **5.7.9** *Calcule los extremos absolutos de* $x^2 + y^2 - xy + x + y$ *sobre el compacto*

$$\{(x, y) \in \mathbb{R}^2 : x \leq 0, y \leq 0, x + y + 3 \geq 0\}$$

♢ **5.7.10** *Obtenga la mínima distancia entre la recta* $y - x + 5 = 0$ *y la parábola* $y = x^2$.

♢ **5.7.11** *Obtenga la distancia de* $(0, 0, 0)$ *a cada una de las superficies*

$$x^2 + y^2 - z^2 + 2xy = 16; \quad z^2 - xy - 1 = 0$$

♢ **5.7.12** *Determine las dimensiones de una caja rectangular sin tapa, con superficie de* 16 m^2*, que encierre un volumen máximo. Justifique la existencia de la caja de volumen máximo.*

♢ **5.7.13** *Calcule el máximo de* $\log x + \log y + 3 \log z$*, sobre la porción de esfera* $x^2+y^2+z^2 = 5r^2$ *en la que* $x > 0$*,* $y > 0$*,* $z > 0$*. Aplíque el resultado para demostrar que si* $a, b, c > 0$ *entonces*

$$abc^3 \leq 27 \left(\frac{a + b + c}{5}\right)^5$$

♢ **5.7.14** *Utilizando la definición compruebe que* $f(x, y, z) = (x - 1)^3 yz + x^2 + 2y^2$ *es diferenciable en* $(1, 0, 0)$ *y obtenga la diferencial* $df(1, 0, 0)$.

♢ **5.7.15** *Sea* $f : (a, b) \to \mathbb{R}$ *derivable en* $x_0 \in (a, b)$*. Demuestre que la función de tres variables* $F(x, y, z) = f(x)$*, definida en* $A = \{(x, y, z) : a < x < b\}$*, es diferenciable en* (x_0, y, z)*. Obtenga* $dF(x_0, y, z)$.

♢ **5.7.16** *Se supone que* $f : \mathbb{R}^n \to \mathbb{R}$ *verifica* $|f(\mathbf{x})| \leq \|\mathbf{x}\|^p$ *para todo* $\mathbf{x} \in \mathbb{R}^n$*, donde* $p > 1$*. Demuestre que* f *es diferenciable en* 0 *y obtenga* $df(\mathbf{0})$.

♢ **5.7.17** *Si* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ *verifica* $\|\mathbf{f}(x) - \mathbf{f}(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|^2$ *para todo* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*, demuestre que* $\mathbf{f}$ *es constante.*

♢ **5.7.18** *Sea* $f : \mathbb{R}^2 \to \mathbb{R}$ *diferenciable en* $(0, 0)$*. Obtenga* $df(0, 0)$ *sabiendo que* $D_{\mathbf{u}} f(0, 0) = 1$ *y* $D_{\mathbf{v}} f(0, 0) = 0$*, donde* $\mathbf{u} = (1, 1)$ *y* $\mathbf{v} = (-1, 1)$.

♢ **5.7.19** *Se consideran las siguientes funciones* $f : \mathbb{R}^2 \to \mathbb{R}$:

a) $f(x, y) = \dfrac{xy^2}{x^2 + y^4}$ *si* $(x, y) \neq (0, 0)$*,* $f(0, 0) = 0$.

b) $f(x, y) = x \operatorname{sen} \dfrac{1}{x^2 + y^2}$ *si* $(x, y) \neq (0, 0)$*,* $f(0, 0) = 0$.

c) $f(x,y) = \dfrac{x|y|}{\sqrt{x^2+y^2}}$ *si* $(x,y) \neq (0,0)$, $f(0,0) = 0$.

d) $f(x,y) = (x^2+y^2)\operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}$ *si* $(x,y) \neq (0,0)$, $f(0,0) = 0$.

e) $f(x.y) = xy^2 \operatorname{sen}(1/y)$ *si* $y \neq 0$, $f(x,0) = 0$.

f) $f(x,y) = (x+y)^n \operatorname{sen}\dfrac{1}{\sqrt{x^2+y^2}}$ *si* $(x,y) \neq (0,0)$, $f(0,0) = 0$.

g) $f(x,y) = \dfrac{x^n}{x^2+y^2}$ *si* $(x,y) \neq (0,0)$, $f(0,0) = 0$. *(*$n \in \mathbb{N}$*)*

Compruebe las afirmaciones que se hacen sobre de cada una de ellas:
a) f *no es diferenciable en* $(0,0)$ *porque no es continua en* $(0,0)$.
b) f *no es diferenciable en* $(0,0)$ *porque no existe una derivada parcial en* $(0,0)$.
c) f *es continua en* $(0,0)$ *y existe la derivada* $D_{\mathbf{u}}f(0,0)$ *según cualquier vector* $\mathbf{u} \in \mathbb{R}^2$ *pero* f *no es diferenciable en* $(0,0)$.
d) f *es diferenciable en* $(0,0)$ *y sus derivadas parciales no son continuas en* $(0,0)$.
e) f *es diferenciable en* $(0,0)$ *porque se cumple la condición suficiente de diferenciabilidad (una de las dos derivadas parciales es continua en* $(0,0)$*).*
f) f *es continua; es diferenciable si y sólo si* $n \geq 2$*; tiene derivadas parciales continuas si y sólo si* $n \geq 3$.
g) f *es diferenciable en* $(0,0)$ *si y sólo si* $n > 3$.

♢ **5.7.20** *Estudie, según los valores de* $n \in \mathbb{N}$*, la continuidad y diferenciabilidad en* $(0,0)$ *de la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por*

$$f(x,y) = \frac{xy^n}{x^2+y^4} \quad \textit{si } (x,y) \neq (0,0); \quad f(0,0) = 0.$$

♢ **5.7.21** *Sea* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por* $f(x,y) = \dfrac{y(x^2-y^2)}{x^2+y^2}$ *si* $(x,y) \neq (0,0)$, $f(0,0) = 0$.
¿Es f *diferenciable en* $(0,0)$*?. ¿Es de clase* $C^1(\mathbb{R}^2)$*?. ¿Es uniformemente continua?.*

♢ **5.7.22** *Sea* $\varphi : \mathbb{R} \to \mathbb{R}$ *una función derivable de la forma* $\varphi(t) = \alpha t + t^2 A(t)$ *donde* $\lim_{t \to 0} A(t) = 1$*. Se define* $f(x,y) = \varphi(xy)/(xy)$ *si* $xy \neq 0$; $f(x,y) = \alpha$ *si* $xy = 0$*. Justifique que* f *es diferenciable en todo punto.*

♢ **5.7.23** *Sea* $f(x,y) = g(x^2+y^2)$ *donde* $g : [0,+\infty) \to \mathbb{R}$ *es derivable dos veces en* $(0,+\infty)$ *y* $g(0) = 0$*. Demuestre las siguientes afirmaciones*

a) *Si existe* $D_1 f(0,0) = A$ *entonces para todo* $\mathbf{u} \in \mathbb{R}^2$ *existe* $D_{\mathbf{u}}f(0,0)$ *y obtenga su valor en función de* A.

b) f *es diferenciable en* $(0,0)$ *si y sólo si existe y es nula la derivada* $D_1 f(0,0)$.

c) f es de clase C^1 si y sólo si y $\lim_{r\to 0} rg'(r^2) = 0$.

♢ **5.7.24** *Determíne los valores de $p > 0$ para los que*

$$f(x,y) = (x^2+y^2)^p \operatorname{sen}\left(\frac{1}{x^2+y^2}\right) \text{ si } (x,y) \neq (0,0), \quad f(0,0) = 0$$

es diferenciable. ¿Para qué valores de p es de clase C^1?.

♢ **5.7.25** *Estudie, según los valores del número real $\alpha > 0$ la diferenciabilidad en $(0,0)$ de la función*

$$f(x,y) = \frac{x}{(x^2+y^2)^\alpha} \operatorname{sen}(\sqrt{x^2+y^2}) \quad \text{si} \quad (x,y) \neq (0,0); \quad f(0,0) = 0$$

♢ **5.7.26** *Sea $f : \mathbb{R} \to \mathbb{R}$ derivable dos veces en $a \in \mathbb{R}$ y $F : \mathbb{R}^2 \to \mathbb{R}$ definida por*

$$F(x,y) = \frac{f(y)-f(x)}{y-x} \quad \text{si} \quad x \neq y, \quad F(x,x) = f'(x)$$

Demuestre que F es diferenciable en (a,a) y obtenga $dF(a,a)$. (Indicación: Considere el desarrollo de Taylor de f en el punto a)

♢ **5.7.27** *Sean $f, g : \Omega \to \mathbb{R}$ definidas en un abierto $\Omega \subset \mathbb{R}^n$.*

a) Si g es continua en $\mathbf{a} \in \Omega$ y f es diferenciable en $\mathbf{a}$ con $f(\mathbf{a}) = 0$, demuestre que fg es diferenciable en $\mathbf{a}$ (aunque g no sea diferenciable en $\mathbf{a}$).

b) Compruebe que es cierto lo que se afirma en el apartado anterior cuando $\Omega = \mathbb{R}^2$, $\mathbf{a} = (0,0)$,

$$f(x,y) = e^x \operatorname{sen} y; \quad g(x,y) = x^3/(x^2+y^2) \quad \text{si} \quad (x,y) \neq (0,0), \ g(0,0) = 0$$

(Estudie la continuidad y diferenciabilidad de g en $(0,0)$).

♢ **5.7.28** *Sean $(E, \|\ \|)$, $(F, \|\ \|)$ espacios normados reales y $\mathbf{f} : E \to F$ una aplicación homogénea: $\mathbf{f}(t\mathbf{x}) = t\mathbf{f}(\mathbf{x})$ para todo $\mathbf{x} \in E$ y todo $t > 0$. Demuestre que $\mathbf{f}$ es diferenciable en 0 si y sólo sí es una aplicación lineal continua.*

♢ **5.7.29** *Demuestre que en un espacio normado $(E, \|\ \|)$ la norma nunca es diferenciable en $\mathbf{0}$. Si la norma procede de un producto escalar, demuestrte que es diferenciable en cada $\mathbf{a} \neq \mathbf{0}$ y obtenga su diferencial.*

♢ **5.7.30** *Sea $(E, \|\ \|)$ un espacio normado real cuya norma procede de un producto escalar $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \mid \mathbf{x} \rangle}$. Si $L : E \to E$ es una aplicación lineal continua, demuestre que $f(\mathbf{x}) = \langle L(\mathbf{x}) \mid \mathbf{x} \rangle$ es diferenciable en todo $\mathbf{a} \in E$ y obtenga su diferencial $df(\mathbf{a})$.*

♢ **5.7.31** *Dadas dos aplicaciones lineales $A, B : \mathbb{R}^n \to \mathbb{R}^m$, se considera la función $f : \mathbb{R}^n \to \mathbb{R}$ definida por $f(\mathbf{x}) = \langle A(\mathbf{x}) | B(\mathbf{x}) \rangle$. Demuestre que f es diferenciable y obtenga su diferencial $df(\mathbf{a})$ en un punto genérico $\mathbf{a} \in \mathbb{R}^n$.*

♢ **5.7.32** *Sea Ω un abierto de un espacio normado $(E, \| \ \|)$ y $\mathbf{f} : \Omega \to E$ una aplicación de la forma $\mathbf{f}(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{x}$ donde $\alpha : \Omega \to \mathbb{R}$ es diferenciable en $\mathbf{a} \in \Omega$. Demuestre que $\mathbf{f}$ también es diferenciable en $\mathbf{a}$ y obtenga una fórmula para $d\mathbf{f}(\mathbf{a})$.*

Aplicación: Si la norma de $(E, \| \ \|)$ procede de un producto escalar, obtenga que $f(\mathbf{x}) = \|\mathbf{x}\| \, \mathbf{x}$ es diferenciable en $\mathbf{a} \in E \setminus \{\mathbf{0}\}$ y calcule $df(\mathbf{a})$.

♢ **5.7.33** *En un espacio normado $(E, \| \ \|)$ la norma procede de un producto escalar. Demuestre que $f(\mathbf{x}) = \dfrac{\mathbf{x}}{\|\mathbf{x}\|}$ es diferenciable en cada $\mathbf{a} \in E \setminus \{\mathbf{0}\}$ y calcule $df(\mathbf{a})$.*

♢ **5.7.34** *Se dice que $f : \mathbb{R}^n \to \mathbb{R}$ es homogénea de grado m si $f(t\mathbf{x}) = t^m f(\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$ y todo $t > 0$. Si f es diferenciable, demuestre que son equivalentes*

i) *f es homogénea de grado m.*

ii) $\langle \nabla f(\mathbf{x}) \mid \mathbf{x} \rangle \ = m f(\mathbf{x})$ *para todo $\mathbf{x} \in \mathbb{R}^n$.*

♢ **5.7.35** *Si $f : \Omega \to \mathbb{R}$ es diferenciable en $\Omega = \mathbb{R}^n \setminus \{\mathbf{0}\}$, demuestre que son equivalentes:*

i) *Existe una función derivable $g : (0, +\infty) \to \mathbb{R}$ tal que $f(\mathbf{x}) = g(\|\mathbf{x}\|_2)$ para todo $\mathbf{x} \in \Omega$.*

ii) *Existe una función $\alpha : \Omega \to \mathbb{R}$ tal que $\nabla f(\mathbf{x}) = \alpha(\mathbf{x})\mathbf{x}$ para todo $\mathbf{x} \in \Omega$.*

Si se cumple i) ¿qué relación hay entre g y α?.

♢ **5.7.36** *Sea Arctg $: \mathbb{R} \to (-\pi/2, \pi/2)$ la rama principal de la función arctg. En el abierto $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$ se define la función $f : \Omega \to \mathbb{R}$*
$f(x,y) = \mathrm{Arctg}(y/x^2)$ *si* $x \neq 0$.
$f(0,y) = -\pi/2$ *si* $y < 0$; $\qquad f(0,y) = \pi/2$ *si* $y > 0$
Demuestre que f es de clase $C^1(\Omega)$ y estudie su continuidad uniforme en Ω.

♢ **5.7.37** *Si $f : \mathbb{R}^n \to \mathbb{R}$ es de clase C^1 y $f(0) = 0$, demuestre que existen funciones continuas, $g_i : \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq n$, tales que $f(\mathbf{x}) = \sum_{i=1}^n x_i g_i(\mathbf{x})$ para todo $\mathbf{x} \in \mathbb{R}^n$.*

♢ **5.7.38** *Demuestre que es de clase C^1 la función*

$$f(x,y) = \frac{x^4 + \operatorname{sen} y^4}{x^2 + y^2} \ \text{ si } (x,y) \neq (0,0), \quad f(0,0) = 0$$

♢ **5.7.39** *Se supone que $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$, es diferenciable en $\mathbf{a}$, que $g : \mathbb{R}^m \to \mathbb{R}$ es diferenciable en $\mathbf{b} = f(\mathbf{a})$ y que $\boldsymbol{\gamma} : [-1,1] \to \mathbb{R}^n$ es derivable en 0. Si $\boldsymbol{\gamma}'(0) = \mathbf{u}$ y $\boldsymbol{\gamma}(0) = \mathbf{a}$ demuestre que $\varphi(t) = g(\mathbf{f}(\boldsymbol{\gamma}(t)))$ es derivable en $t = 0$ y que $\varphi'(0) = \langle \nabla g(\mathbf{b}) | D_{\mathbf{u}}\mathbf{f}(\mathbf{a}) \rangle$*

♢ **5.7.40** *Escriba, en un punto genérico, la matriz jacobiana de las aplicaciones*

a) $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$, $\ \mathbf{f}(x,y) = (\operatorname{sen}(xy), \operatorname{sen}(x \operatorname{sen} y), x^4)$;

b) $\mathbf{F} = f \circ \mathbf{g}$ *donde* $f(x, y, z) = x^2 + y^2 + z^2$, $\mathbf{g}(u, v) = (u - v, u - v, u^2 - v^2)$.

♢ **5.7.41** *Sea* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$, *la aplicación definida por*

$$\mathbf{f}(x, y) = (x + \cos y, \ \frac{x^4}{x^2 + y^2}) \ \ \textit{si } (x, y) \neq (0, 0), \ \mathbf{f}(0, 0) = (1, 0).$$

Justifique que $\mathbf{F} = \mathbf{f} \circ \mathbf{f} + \mathbf{f}$ *es diferenciable en* $(0, 0)$ *y calcule* $d\mathbf{F}(0, 0)$.

♢ **5.7.42** *Sean* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$, $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ *definidas por*

$$\mathbf{f}(x, y) = (xy, \ e^x \cos y, \ e^{xy}), \quad \mathbf{g}(u, v) = (u + e^v, \ u^3 v^2)$$

Obtenga que $\mathbf{F} = \mathbf{f} \circ \mathbf{g}$ *es diferenciable en todo punto y calcule* $d\mathbf{F}(0, 0)$.

♢ **5.7.43** *¿En qué puntos es diferenciable la función* $f(x, y) = \sqrt{1 + |xy|}$*?. Demuestre que* $|D_{\mathbf{u}} f(1, -1)| \leq \|\mathbf{u}\|_2 / 2$ *para todo* $\mathbf{u} \in \mathbb{R}^2$.

♢ **5.7.44** *Si* $f(x, y, z) = x^2 - y^2 + xyz^2 - zx$, *obtenga el máximo valor de las derivadas direccionales* $\{D_{\mathbf{u}} f(1, 2, 3) : \|\mathbf{u}\|_2 = 1\}$

♢ **5.7.45** *Obtenga* $a, b, c \in \mathbb{R}$ *con la condición de que la derivada direccional de* $f(x, y, z) = axy^2 + byz + cz^2x^3$ *en el punto* $(1, 2, -1)$ *alcance un valor máximo, igual a* 64, *según la dirección del vector* $(0, 0, 1)$.

♢ **5.7.46** *Sea* $f(x, y) = \dfrac{x^3 + 2xy^2}{x^2 + y^2}$ *si* $(x, y) \neq (0, 0)$, $f(0, 0) = 0$
Calcule el máximo valor de la derivada direccional $D_{\mathbf{u}} f(0, 0)$, $\|\mathbf{u}\|_2 = 1$.

♢ **5.7.47** *Sea* $\Omega \subset \mathbb{R}^n$ *abierto y* $\mathbf{a} \in \Omega$. *Si* $f : \Omega \to \mathbb{R}$ *es continua en* $\mathbf{a}$, *diferenciable en cada* $\mathbf{x} \in \Omega \setminus \{\mathbf{a}\}$, *y existen los límites* $\lim_{\mathbf{x} \to \mathbf{a}} D_k f(\mathbf{x}) = A_k$, $1 \leq k \leq n$, *demuestre que* f *es diferenciable en* $\mathbf{a}$.
(Indic. Considere $\varphi(t) = f(a + th) - t \sum_{k=1} A_k h_k$*)*

♢ **5.7.48** *Sean* $(E, \| \ \|)$, $(F, \| \ \|)$ *espacios normados,* $\mathbf{f} : \Omega \to F$ *una función continua, definida en un abierto* $\Omega \subset E$ *y* $\mathbf{a} \in \Omega$. *Si* $\mathbf{f}$ *diferenciable en cada* $\mathbf{x} \in \Omega \setminus \{\mathbf{a}\}$, *y existe* $\lim_{\mathbf{x} \to \mathbf{a}} d\mathbf{f}(\mathbf{x}) = L$, *demuestre que* $\mathbf{f}$ *es diferenciable en* $\mathbf{a}$ *y que* $d\mathbf{f}(\mathbf{a}) = L$.

♢ **5.7.49** *Sean* $(E, \| \ \|)$, $(F, \| \ \|)$ *espacios normados y* $\mathbf{f} : \Omega \to F$ *una función continua definida en un abierto* $\Omega \subset E$. *Se supone que para cada* $\mathbf{x} \in \Omega$ *y cada* $\mathbf{u} \in E$ *existen las derivadas* $D_{\mathbf{u}} \mathbf{f}(\mathbf{x}) = A(\mathbf{x})\mathbf{u}$ *donde* $A : \Omega \to \mathcal{L}(E, F)$ *es continua. Demuestre que* f *es diferenciable y que* $d\mathbf{f}(\mathbf{x}) = A(\mathbf{x})$ *para todo* $\mathbf{x} \in \Omega$.

♢ **5.7.50** *Si* $\Omega \subset \mathbb{R}^n$ *es un abierto convexo y* $f : \Omega \to \mathbb{R}$ *es diferenciable con derivadas parciales acotadas demuestre que* f *es uniformemente continua.*

♢ **5.7.51** *Sea* $\Omega \subset \mathbb{R}^n$ *abierto convexo y* $(E, \| \ \|)$ *un espacio normado completo.*
Si $\mathbf{f} : \Omega \to (E, \| \ \|)$ *es de clase* $C^1(\Omega, E)$ *y* $\|d\mathbf{f}(\mathbf{x})\| \leq M$ *para todo* $\mathbf{x} \in \Omega$, *demuestre que* $\mathbf{f}$ *admite una extensión continua* $\hat{\mathbf{f}} : \overline{\Omega} \to E$.

♢ **5.7.52** *Escriba las ecuaciones de los planos tangentes a las siguientes superficies en los puntos que se indican:*

a) Superficie de ecuación $x^2 + y^2 + 4z^2 = 12$*, en el punto* $\mathbf{p} = (2, 2, 1)$*;*

b) Superficie de ecuaciones paramétricas

$$x(r, t) = r \cos t, \quad y(r, t) = r \operatorname{sen} t, \quad z(r, t) = t, \quad 0 < r, \quad 0 < t < 2\pi$$

en el punto $\mathbf{p} = (0, 2, \pi/2)$ *imagen de* $(2, \pi/2)$.

♢ **5.7.53** *Se considera la función* $f : \mathbb{R}^3 \to \mathbb{R}$ *definida por*

$$f(x, y, z) = \frac{x}{\sqrt{x^2 + y^2}} \operatorname{sen}(x^2 + y^2) + e^z \text{ si } (x, y, z) \neq (0, 0, z), \quad f(0, 0, z) = e^z$$

Compruebe que f *es diferenciable en* $(0, 0, 0)$ *y obtenga la derivada* $D_{\mathbf{v}} f(0, 0, 0)$ *según un vector unitario* $\mathbf{v}$ *normal a la superficie* $x + y^2 + 2z = 0$ *en el punto* $(0, 0, 0)$.

♢ **5.7.54** *Demuestre que la función* $f(x, y, z) = \int_y^z e^{-x^2t^2} dt$ *es diferenciable en todos los puntos. Calcule la derivada* $D_{\mathbf{v}} f(\mathbf{p})$ *donde* $\mathbf{p} = (2, 2, 1)$ *y* $\mathbf{v}$ *es un vector unitario normal al elipsoide* $2x^2 + y^2 + z^2 = 13$*, en el punto* $\mathbf{p}$*, que apunta hacia el exterior del elipsoide.*

♢ **5.7.55** *Sea* $f : \mathbb{R}^3 \to \mathbb{R}$ *definida por*

$$f(x, y, z) = \frac{x^2 z^2}{x^2 + y^2} \quad \text{si} \quad (x, y) \neq (0, 0), \quad f(0, 0, z) = 0.$$

Obtenga $D_{\mathbf{v}} f(1, 1, \sqrt{2})$*, donde* $\mathbf{v}$ *es un vector unitario tangente a la curva*

$$x^2 + y^2 = 2x, \quad x^2 + y^2 + z^2 = 4$$

en el punto $(1, 1, \sqrt{2})$.

♢ **5.7.56** *Sea* $f(x, y) = y^2 \operatorname{sen}(x/y)$ *si* $y \neq 0$*,* $f(x, 0) = 0$.

i) Estudie la diferenciabilidad de f *en un punto genérico* $(a, b) \in \mathbb{R}^2$

ii) Estudie la continuidad y diferenciabilidad de $D_1 f(x, y)$ *y* $D_2 f(x, y)$ *en el punto* $(0, 0)$.

iii) Determine un vector unitario $u \in \mathbb{R}^2$ *tal que* $D_u f(\pi, 1) = 1$.

♢ **5.7.57** *Estudie la diferenciabilidad en* $(0, 0, 0)$ *de la función* $f : \mathbb{R}^3 \to \mathbb{R}$ *definida por*

$$f(x, y, z) = z + y \operatorname{sen} \frac{x^2}{\sqrt{x^2 + y^2}} \quad \text{si } (x, y) \neq (0, 0), \quad f(0, 0, z) = z$$

Dado un punto $\mathbf{p}$ *del elipsoide* $E = \{(x, y, z) : x^2 + 2y^2 + 2z^2 = 2\}$*, obtenga un vector unitario* $\mathbf{v}$ *tangente a* E *en* $\mathbf{p}$ *que proporcione el mayor valor de la derivada* $D_{\mathbf{v}} f(\mathbf{p})$.