

Matriz jacobiana

Como último caso particular de la noción de diferenciabilidad, suponemos ahora que el espacio normado de partida es $\mathbb{R}^N$ con $N > 1$, y el de llegada es $\mathbb{R}^M$, también con $M > 1$. Estudiamos por tanto la diferenciabilidad de una función definida en un abierto de $\mathbb{R}^N$ y con valores en $\mathbb{R}^M$, es decir, de un campo vectorial. Dependiendo de los valores de N y M tenemos campos vectoriales muy variados, siendo lógicamente $N = M = 2$ y $N = M = 3$, los casos más interesantes. Usando las bases usuales de $\mathbb{R}^N$ y $\mathbb{R}^M$, la diferencial de nuestra función en un punto, cuando existe, viene representada por una matriz $M \times N$ con coeficientes reales, que será la *matriz jacobiana*. Sus filas son los gradientes de M campos escalares en $\mathbb{R}^N$, las M componentes de nuestro campo vectorial.

A la composición de aplicaciones lineales corresponde entonces el producto de matrices, con lo que obtenemos una *regla de la cadena para las derivadas parciales*. Como aplicación geométrica que merece destacarse, cuando $N = 2$ y $M = 3$ estudiamos el plano tangente a una *superficie paramétrica* en $\mathbb{R}^3$, generalizando lo que ya sabemos para superficies explícitas.

9.1. Matriz de una aplicación lineal

Toda aplicación lineal $T \in L(\mathbb{R}^N, \mathbb{R}^M)$ admite una expresión matricial que vamos a recordar. Denotamos por $\mathcal{M}_{M \times N}$ al conjunto de todas las matrices $M \times N$ (M filas y N columnas) con coeficientes reales, cuya estructura algebraica suponemos bien conocida. Para cada $x \in \mathbb{R}^N$, sus coordenadas en la base usual de $\mathbb{R}^N$ forman una matriz columna que también denotamos por x , por lo que escribimos $x \in \mathcal{M}_{N \times 1}$. Análogamente, el vector $y = T(x) \in \mathbb{R}^M$ tiene sus coordenadas en la base usual de $\mathbb{R}^M$, que dan la matriz columna $y \in \mathcal{M}_{M \times 1}$. Existe entonces una única matriz $A \in \mathcal{M}_{M \times N}$ tal que, para todo $x \in \mathbb{R}^N$, obtenemos $y = T(x)$ como producto de matrices: $y = A \cdot x$.

De hecho, si escribimos $A = (\alpha_{jk})$ para indicar los coeficientes de la matriz A , se tiene

$$\alpha_{jk} = (\pi_j \circ T)(e_k) \quad \forall j \in \Delta_M, \quad \forall k \in \Delta_N \quad (1)$$

siendo $\{e_1, \dots, e_N\}$ la base usual de $\mathbb{R}^N$ y $\{\pi_1, \dots, \pi_M\}$ las proyecciones coordenadas en $\mathbb{R}^M$.

Diremos que A es la **matriz de la aplicación lineal** T , pero debemos entender su unicidad: es la única matriz que representa a T en la forma $y = A \cdot x$ cuando, tanto para $x \in \mathbb{R}^N$ como para $y = T(x) \in \mathbb{R}^M$, usamos sus coordenadas en las bases usuales, escritas en forma de matrices columna. Tenemos de hecho un isomorfismo entre los espacios vectoriales $L(\mathbb{R}^N, \mathbb{R}^M)$ y $\mathcal{M}_{M \times N}$, que identifica cada $T \in L(\mathbb{R}^N, \mathbb{R}^M)$ con su matriz $A \in \mathcal{M}_{M \times N}$, recién definida.

9.2. Matriz jacobiana

En lo que sigue fijamos un abierto Ω de $\mathbb{R}^N$ y una función $f : \Omega \rightarrow \mathbb{R}^M$. Cuando $M = 1$ tenemos un campo escalar en $\mathbb{R}^N$, mientras que cuando $N = 1$ se trata de una función de variable real con valores en $\mathbb{R}^M$, los casos ya estudiados. Lo que hagamos es válido en ambos casos, pero nos interesa lo que ocurre para $N > 1$ y $M > 1$. Escribimos $f = (f_1, f_2, \dots, f_M)$ para indicar las M componentes de f , campos escalares definidos en Ω . Recordamos que, para todo $j \in \Delta_M$ se tiene $f_j = \pi_j \circ f$.

Si f es diferenciable en un punto $a \in \Omega$, la matriz de la aplicación lineal $Df(a) \in L(\mathbb{R}^N, \mathbb{R}^M)$ recibe el nombre de **matriz jacobiana** de f en a y se denota por $Jf(a)$. Por tanto, la matriz jacobiana representa a la diferencial $Df(a)$ en el siguiente sentido: para todo $x \in \mathbb{R}^N = \mathcal{M}_{N \times 1}$, el vector $y = Df(a)(x) \in \mathbb{R}^M = \mathcal{M}_{M \times 1}$ se obtiene como producto de matrices: $y = Jf(a) \cdot x$.

De (1) deducimos que los coeficientes de la matriz jacobiana, $Jf(a) = (\alpha_{jk}) \in \mathcal{M}_{M \times N}$, vienen dados por $\alpha_{jk} = (\pi_j \circ Df(a))(e_k)$, para cualesquiera $j \in \Delta_M$ y $k \in \Delta_N$. Fijado $j \in \Delta_M$, sabemos que $f_j = \pi_j \circ f$ es diferenciable en a con $Df_j(a) = \pi_j \circ Df(a)$. Pero además, para cada $k \in \Delta_N$, $Df_j(a)(e_k)$ es la derivada parcial de f_j con respecto a la k -ésima variable en el punto a , luego

$$\alpha_{jk} = Df_j(a)(e_k) = \frac{\partial f_j}{\partial x_k}(a) \quad \forall j \in \Delta_M, \quad \forall k \in \Delta_N$$

Así pues, la matriz jacobiana de f en a viene dada por:

$$Jf(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \dots & \frac{\partial f_1}{\partial x_N}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \dots & \frac{\partial f_2}{\partial x_N}(a) \\ \vdots & \vdots & & & \vdots \\ \frac{\partial f_M}{\partial x_1}(a) & \frac{\partial f_M}{\partial x_2}(a) & \dots & \dots & \frac{\partial f_M}{\partial x_N}(a) \end{pmatrix}$$

Se recuerda fácilmente, pues para cada $j \in \Delta_M$, su j -ésima fila es el gradiente de f_j en a , es decir, las M filas de la matriz jacobiana son los gradientes de las M componentes de f . Por otra parte, para cada $k \in \Delta_N$, la k -ésima columna de la matriz jacobiana es el vector derivada parcial de f con respecto a la k -ésima variable en el punto a , luego las N columnas de la matriz jacobiana son las N derivadas parciales de f en a .

Cuando $M = 1$, tenemos un campo escalar, cuya matriz jacobiana es su gradiente, escrito como una matriz fila, $Jf(a) = \nabla f(a) \in \mathcal{M}_{1 \times N}$, de forma que, para todo $x \in \mathbb{R}^N$, el producto escalar $(\nabla f(a) | x)$ coincide con el producto de matrices $\nabla f(a) \cdot x = Jf(a) \cdot x$. Por otra parte, cuando $N = 1$, tenemos una función de una sola variable real, cuya matriz jacobiana es su vector derivada, escrito como una matriz columna, $Jf(a) = f'(a) \in \mathcal{M}_{M \times 1}$, de forma que, para todo $x \in \mathbb{R}$, el producto del vector $f'(a)$ por el escalar x coincide con el producto de matrices $f'(a) \cdot x = Jf(a) \cdot x$.

9.3. Regla de la cadena para las derivadas parciales

La noción de matriz jacobiana, junto con la regla general de la cadena, nos van a permitir calcular las derivadas parciales de las componentes de una composición de campos vectoriales diferenciables.

Mantenemos pues nuestra función $f : \Omega \rightarrow \mathbb{R}^M$, donde Ω es un abierto de $\mathbb{R}^N$, pero ahora consideramos un abierto U de $\mathbb{R}^M$ tal que $f(\Omega) \subset U$ y otra función $g : U \rightarrow \mathbb{R}^P$ donde $P \in \mathbb{N}$ es arbitrario. Esto permite definir la composición $h = g \circ f : \Omega \rightarrow \mathbb{R}^P$. Suponemos que f es diferenciable en un punto $a \in \Omega$ y que g es diferenciable en el punto $b = f(a) \in U$, con lo que la regla de la cadena nos dice que h es diferenciable en a con

$$Dh(a) = Dg(b) \circ Df(a)$$

Veamos cómo se traduce esta igualdad en términos de las matrices jacobianas. Para $x \in \mathbb{R}^N$, sean $y = Df(a)(x) \in \mathbb{R}^M$ y $z = Dg(b)(y) \in \mathbb{R}^P$, con lo que $z = Dh(a)(x)$. Escribiendo los tres vectores x, y, z como matrices columna, tenemos

$$Jh(a) \cdot x = z = Jg(b) \cdot y = Jg(b) \cdot (Jf(a) \cdot x) = (Jg(b) \cdot Jf(a)) \cdot x \quad \forall x \in \mathbb{R}^N$$

donde hemos usado la asociatividad del producto de matrices. Deducimos claramente que

$$Jh(a) = Jg(b) \cdot Jf(a) \quad (2)$$

Esta igualdad matricial nos dará una regla práctica para calcular derivadas parciales, tan pronto como especifiquemos los coeficientes de las tres matrices que en ella aparecen. Para $Jf(a)$ ya lo hemos hecho antes:

$$Jf(a) = (\alpha_{jk}) = \left(\frac{\partial f_j}{\partial x_k}(a) \right) \in \mathcal{M}_{M \times N}$$

La función g tiene P componentes g_i con $i \in \Delta_P$, que son funciones de M variables reales. A efectos de escribir las derivadas parciales de cada componente, la j -ésima de esas variables debe denotarse lógicamente por y_j para todo $j \in \Delta_M$. De esta forma, para $i \in \Delta_P$ y $j \in \Delta_M$, la derivada parcial de g_i con respecto a la j -ésima variable, en el punto b , es $\frac{\partial g_i}{\partial y_j}(b)$, y ya podemos escribir explícitamente la matriz jacobiana:

$$Jg(b) = (\beta_{ij}) = \left(\frac{\partial g_i}{\partial y_j}(b) \right) \in \mathcal{M}_{P \times M}$$

Finalmente h tiene también sus P componentes $h_i = g_i \circ f$ con $i \in \Delta_P$, que son funciones de N variables reales, las mismas de las que depende f , esto es, x_k con $k \in \Delta_N$. Por tanto:

$$Jh(a) = (\lambda_{ik}) = \left(\frac{\partial h_i}{\partial x_k}(a) \right) \in \mathcal{M}_{P \times N}$$

En vista de (2) la definición del producto de matrices nos dice que, para cualesquiera $i \in \Delta_P$ y $k \in \Delta_N$, se tiene $\lambda_{ik} = \sum_{j=1}^M \beta_{ij} \alpha_{jk}$, es decir,

$$\frac{\partial h_i}{\partial x_k}(a) = \sum_{j=1}^M \frac{\partial g_i}{\partial y_j}(b) \frac{\partial f_j}{\partial x_k}(a) \quad \forall k \in \Delta_N, \quad \forall i \in \Delta_P \quad (3)$$

Podemos así calcular las derivadas parciales de todas las componentes de h , una vez que hayamos hecho lo mismo con f y g . También podemos usar (3) como sistema de ecuaciones, del que obtener ciertas derivadas parciales a partir de otras que conozcamos. Enseguida veremos ejemplos de aplicación de esta regla práctica para el cálculo de derivadas parciales.

9.4. Cambio de variables

La igualdad (3) será más fácil de entender, y de recordar, con una notación más intuitiva, la que se usa en la práctica para manejar diversos ejemplos de cambio de variables. Mantenemos las hipótesis que nos han permitido obtener (3), es decir, f es diferenciable en $a \in \Omega$ y g es diferenciable en $b = f(a) \in U$.

Pensemos que f , g y h describen la relación entre tres variables, $x \in \Omega$, $y \in U$ y $z \in \mathbb{R}^P$, mediante las tres igualdades

$$y = f(x), \quad z = g(y), \quad z = h(x)$$

siendo la tercera consecuencia de las dos primeras. Nótese el doble papel de la variable y , que depende de x mediante la función f , pero es también la variable de la que depende z mediante la función g . Como consecuencia z también depende de x mediante la función h . Nada de particular, es lo que ocurre siempre que consideramos una composición de funciones, entendidas como dependencia entre variables. La función $z = h(x)$ para $x \in \Omega$, se ha obtenido a partir de la función $z = g(y)$ para $y \in U$, mediante el cambio de variable $y = f(x)$ para $x \in \Omega$.

Si ahora usamos las coordenadas en las bases usuales de los tres espacios involucrados, las tres variables vectoriales x, y, z se convierten en tres bloques de variables que ya toman valores reales: $x = (x_1, x_2, \dots, x_N)$, $y = (y_1, y_2, \dots, y_M)$ y $z = (z_1, z_2, \dots, z_P)$. Estas $N + M + P$ variables están relacionadas mediante las componentes de f , g y h , por las igualdades

$$y_j = f_j(x_1, x_2, \dots, x_N), \quad z_i = g_i(y_1, y_2, \dots, y_M), \quad z_i = h_i(x_1, x_2, \dots, x_N) \quad (4)$$

válidas para cualesquiera $j \in \Delta_M$, $i \in \Delta_P$. El doble papel de las variables vectoriales y, z se ha transmitido a las variables reales $y_1, y_2, \dots, y_M$ y $z_1, z_2, \dots, z_P$.

Las igualdades (4) sugieren sustituir en (3) cada función por la variable cuyos valores determina, teniendo en cuenta que g_i y h_i determinan a la misma variable z_i , la primera como función de $y_1, y_2, \dots, y_M$ y la segunda como función de $x_1, x_2, \dots, x_N$. Así pues, escribimos

$$\frac{\partial y_j}{\partial x_k}(a) = \frac{\partial f_j}{\partial x_k}(a), \quad \frac{\partial z_i}{\partial y_j}(b) = \frac{\partial g_i}{\partial y_j}(b), \quad \frac{\partial z_i}{\partial x_k}(a) = \frac{\partial h_i}{\partial x_k}(a) \quad (5)$$

para cualesquiera $k \in \Delta_N$, $j \in \Delta_M$, $i \in \Delta_P$.

El doble uso de la variable y_j no tiene problema: en la primera igualdad está claro que y_j es una función de $(x_1, x_2, \dots, x_N)$, la función f_j que teníamos en (4), mientras que en la segunda igualdad y_j es una de las variables de las que depende z_i . El doble uso de la variable z_i es más delicado, pero también se comprende fácilmente. La dependencia de z_i con respecto a y_j viene determinada por la función g_i , como puede verse en (4), luego la derivada parcial de z_i con respecto a y_j no puede ser otra que la de g_i . Por la misma razón, en vista de la tercera igualdad de (4), la derivada parcial de z_i con respecto a x_k sólo puede ser la de h_i . En resumidas cuentas, lo que estamos haciendo es entender las derivadas parciales de unas variables respecto de otras como las derivadas parciales de las funciones que determinan la dependencia de las unas respecto de las otras. Esta notación parece complicada, e incluso formalmente discutible, pero es la más intuitiva y cómoda en la práctica, como enseguida veremos.

Así pues, al sustituir en (3) las igualdades (5), la regla de la cadena para el cálculo de derivadas parciales queda de la siguiente forma:

$$\frac{\partial z_i}{\partial x_k}(a) = \sum_{j=1}^M \frac{\partial z_i}{\partial y_j}(b) \frac{\partial y_j}{\partial x_k}(a) \quad \forall k \in \Delta_N \quad \forall i \in \Delta_P \quad (6)$$

De entrada esta igualdad es mucho más fácil de recordar que (3), sólo incluye los tres bloques inevitables de variables, $x \in \Omega$, $y \in U$ y $z \in \mathbb{R}^P$, formalmente han desaparecido las componentes de f, g y h , aunque obviamente siguen estando ahí, son las funciones con cuyas derivadas parciales estamos operando.

Además, (6) tiene una interpretación intuitiva que conviene comentar. Para $i \in \Delta_P$ y $k \in \Delta_N$, la derivada parcial $\frac{\partial z_i}{\partial x_k}(a)$ se obtiene pensando que z_i sólo depende de x_k , cuando el resto de variables x_h con $h \in \Delta_N \setminus \{k\}$ se mantienen constantes, dependencia que se produce, por así decirlo, a través de todas las variables intermedias y_j con $j \in \Delta_M$, pues todas ellas dependen de x_k . Esto permite interpretar la suma del segundo miembro de (6). Si fuese $M = 1$ y sólo hubiese una variable intermedia y_j , la regla de la cadena para dos funciones de una variable nos daría la derivada parcial como el producto $\frac{\partial z_i}{\partial y_j}(b) \frac{\partial y_j}{\partial x_k}(a)$, que es el j -ésimo sumando del segundo miembro de (6), así que podríamos decir que dicho sumando es la contribución de la variable y_j a la derivada parcial que estamos calculando. Entonces (6) nos dice que dicha derivada parcial se obtiene sumando las contribuciones de todas las variables intermedias. Por supuesto, esta interpretación no sirve como demostración de (6), pero ayuda a entender su significado. No obstante, la mejor forma de familiarizarse con la regla obtenida para el cálculo de derivadas parciales es ver cómo se usa en ejemplos concretos.

9.5. Coordenadas polares en el plano

Consideremos el caso particular en que $M = N = 2$ y f es la función que se usa para definir las coordenadas polares, pongamos por caso, en el semiplano superior. Concretamente estamos tomando $\Omega = \mathbb{R}^+ \times]0, \pi[$, y $f : \Omega \rightarrow \mathbb{R}^2$ es la función cuyas componentes, que vamos a denotar por x e y , vienen dadas por

$$x(\rho, \theta) = \rho \cos \theta \quad \text{e} \quad y(\rho, \theta) = \rho \sin \theta \quad \forall (\rho, \theta) \in \Omega \quad (7)$$

Es fácil ver que x e y , luego también f , son diferenciables en todo punto $(\rho, \theta) \in \Omega$, con

$$\frac{\partial x}{\partial \rho}(\rho, \theta) = \cos \theta, \quad \frac{\partial x}{\partial \theta}(\rho, \theta) = -\rho \sin \theta, \quad \frac{\partial y}{\partial \rho}(\rho, \theta) = \sin \theta, \quad \frac{\partial y}{\partial \theta}(\rho, \theta) = \rho \cos \theta \quad (8)$$

Obsérvese lo que ha ocurrido en nuestro caso con la primera igualdad de (5). Las variables de partida x_1 y x_2 son ρ y θ , mientras que las variables intermedias y_1 e y_2 serán x e y . Entonces la primera igualdad de (5) es automática, pues empezamos tomando $f_1 = x$ y $f_2 = y$. Ahora x e y pasarán a ser variables de las que dependerá otra función g , definida en $f(\Omega)$. Se comprueba sin dificultad que $f(\Omega)$ es el semiplano superior.

Consideremos un campo escalar $g : U \rightarrow \mathbb{R}$, donde $U = \{(x, y) \in \mathbb{R}^2 : y > 0\} = f(\Omega)$, y supongamos que g es diferenciable en todo punto de U . Por ejemplo, g puede describir la temperatura en el semiplano superior. Ahora tenemos $P = 1$, luego una sola variable $z \in \mathbb{R}$, que depende de x e y mediante la igualdad $z = g(x, y)$ y nos da la temperatura en el punto cuyas coordenadas cartesianas son (x, y) . Por tanto, para todo $(x, y) \in U$ escribimos

$$\frac{\partial z}{\partial x}(x, y) = \frac{\partial g}{\partial x}(x, y) \quad \text{y} \quad \frac{\partial z}{\partial y}(x, y) = \frac{\partial g}{\partial y}(x, y)$$

Tenemos aquí, en nuestro caso, la segunda igualdad de (5). La notación sugiere claramente que las funciones $\frac{\partial z}{\partial x}$ y $\frac{\partial z}{\partial y}$, las derivadas parciales de g , nos informan sobre cómo varía la temperatura al desplazarnos horizontal o verticalmente desde un punto arbitrario de U .

Para estudiar la temperatura en coordenadas polares, consideramos la función $h = g \circ f$, que es diferenciable en todo punto de Ω , y escribimos

$$z = g(x, y) = g(\rho \cos \theta, \rho \sin \theta) = h(\rho, \theta) \quad \forall (\rho, \theta) \in \Omega$$

Así pues, en nuestro caso, la tercera igualdad de (5) es la siguiente:

$$\frac{\partial z}{\partial \rho}(\rho, \theta) = \frac{\partial h}{\partial \rho}(\rho, \theta) \quad \text{y} \quad \frac{\partial z}{\partial \theta}(\rho, \theta) = \frac{\partial h}{\partial \theta}(\rho, \theta) \quad \forall (\rho, \theta) \in \Omega$$

Obsérvese que h es la función que describe la temperatura en coordenadas polares, es decir, para cada $(\rho, \theta) \in \Omega$, $z = h(\rho, \theta)$ es la temperatura en el punto $a = (x, y) = (\rho \cos \theta, \rho \sin \theta)$, el punto del semiplano superior cuyas coordenadas polares son ρ y θ . El doble uso de z tiene perfecto sentido: cada valor de z es la temperatura en un punto a del semiplano superior, que podemos calcular usando las coordenadas cartesianas de a , mediante la igualdad $z = g(x, y)$, o usando sus coordenadas polares, mediante $z = h(\rho, \theta)$.

Las derivadas parciales de h , es decir, las funciones $\frac{\partial z}{\partial \rho}$ y $\frac{\partial z}{\partial \theta}$, nos informan sobre como varía la temperatura al desplazarnos desde un punto arbitrario de U , por la recta que le une con el origen, o por una circunferencia centrada en el origen.

Pues bien, para $(\rho, \theta) \in \Omega$ arbitrario, la regla de la cadena para las derivadas parciales, escrita como en (6), con $(x, y) = (\rho \cos \theta, \rho \sin \theta)$, nos da:

$$\begin{aligned}\frac{\partial z}{\partial \rho}(\rho, \theta) &= \cos \theta \frac{\partial z}{\partial x}(x, y) + \sin \theta \frac{\partial z}{\partial y}(x, y) & y \\ \frac{\partial z}{\partial \theta}(\rho, \theta) &= -\rho \sin \theta \frac{\partial z}{\partial x}(x, y) + \rho \cos \theta \frac{\partial z}{\partial y}(x, y)\end{aligned}\quad (9)$$

Hemos hecho una serie de comentarios para poner de manifiesto el aspecto que toman en este caso concreto todas las igualdades que habíamos manejado en general, pero en la práctica el razonamiento es mucho más directo y sencillo: partiendo de (7) que no es más que la definición de las coordenadas polares, deducimos claramente (8), y entonces, viendo z como función de las variables x, y , definida en U , y también como función de ρ, θ , definida en Ω , tenemos (9).

Naturalmente, cuando conocemos explícitamente g , también tenemos explícitamente h , y es probable que podamos calcular sus derivadas parciales sin usar la regla de la cadena. Lo interesante de (9) es que permite calcular las derivadas parciales de h a partir de las de g , sin necesidad de conocer explícitamente g . Pero recíprocamente, si conocemos las derivadas parciales de h , podemos calcular las de g , pues de (9) se deduce claramente que

$$\begin{aligned}\cos \theta \frac{\partial z}{\partial \rho}(\rho, \theta) - \frac{\sin \theta}{\rho} \frac{\partial z}{\partial \theta}(\rho, \theta) &= \frac{\partial z}{\partial x}(x, y) & y \\ \sin \theta \frac{\partial z}{\partial \rho}(\rho, \theta) + \frac{\cos \theta}{\rho} \frac{\partial z}{\partial \theta}(\rho, \theta) &= \frac{\partial z}{\partial y}(x, y)\end{aligned}$$

9.6. Relación entre plano tangente y rectas tangentes

La regla de la cadena para las derivadas parciales nos va a dar una propiedad del plano tangente a una superficie explícita, que luego usaremos para motivar la definición del plano tangente a superficies más generales. Consideremos una superficie explícita $\Sigma = \text{Gr } f \subset \mathbb{R}^3$, donde $f: \Omega \rightarrow \mathbb{R}$ es una función continua en un abierto conexo $\Omega \subset \mathbb{R}^2$. Sabemos que si f es diferenciable en un punto $(x_0, y_0) \in \Omega$, tomando $z_0 = f(x_0, y_0)$ y $P_0 = (x_0, y_0, z_0) \in \Sigma$, el plano tangente Π , a la superficie Σ en el punto P_0 viene dado por la ecuación

$$z - z_0 = (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) \quad (10)$$

Sea ahora $C = \gamma(J)$ una curva paramétrica en $\mathbb{R}^3$, tal que $P_0 \in C \subset \Sigma$. Se tiene por tanto que $\gamma: J \rightarrow \mathbb{R}^3$ es continua en un intervalo abierto $J \subset \mathbb{R}$, existe $t_0 \in J$ tal que $\gamma(t_0) = P_0$, y además, $\gamma(t) \in \Sigma$ para todo $t \in J$.

Si x, y, z son las tres componentes de γ , vemos que $(x(t), y(t)) \in \Omega$ y $z(t) = f(x(t), y(t))$ para todo $t \in J$. Así pues, las ecuaciones paramétricas de C son

$$x = x(t), \quad y = y(t), \quad z = f(x(t), y(t)) \quad \forall t \in J$$

Supongamos que $P_0 = \gamma(t_0)$ es punto regular de $C = \gamma(J)$, es decir, que γ es derivable en t_0 con $\gamma'(t_0) \neq 0$, para ver la relación entre la recta R , tangente a la curva C en el punto P_0 , y el plano tangente Π . Podemos entender la recta R como una *recta tangente* a la superficie Σ en el punto P_0 , y cabe esperar que se tenga $R \subset \Pi$, como efectivamente vamos a comprobar.

La regla de la cadena nos dice que

$$z'(t_0) = x'(t_0) \frac{\partial f}{\partial x}(x_0, y_0) + y'(t_0) \frac{\partial f}{\partial y}(x_0, y_0)$$

luego las ecuaciones paramétricas de R son:

$$\begin{aligned} x &= x_0 + t x'(t_0), & y &= y_0 + t y'(t_0) & y \\ z &= z_0 + t \left(x'(t_0) \frac{\partial f}{\partial x}(x_0, y_0) + y'(t_0) \frac{\partial f}{\partial y}(x_0, y_0) \right) & \forall t \in \mathbb{R} \end{aligned} \quad (11)$$

Comparando (11) y (10) vemos claramente que para todo $(x, y, z) \in R$ se tiene $(x, y, z) \in \Pi$, luego $R \subset \Pi$ como queríamos.

En resumen, podemos decir que *el plano tangente a la superficie Σ en el punto P_0 contiene a todas las rectas tangentes en dicho punto a curvas paramétricas que estén contenidas en la superficie Σ* , o si se quiere, a todas las rectas tangentes a la superficie Σ en P_0 . La definición del plano tangente se justificó por ser una buena aproximación de la superficie Σ cerca del punto P_0 , pero ahora hemos encontrado una segunda justificación, más intuitiva si cabe: el plano tangente a Σ en P_0 es el único plano que contiene a todas las rectas tangentes a Σ en P_0 .

9.7. Superficies en forma paramétrica

Generalizando la noción de superficie explícita, definimos las superficies paramétricas, de forma bastante análoga a lo que hicimos para las curvas paramétricas, sólo que intuitivamente ahora se trata de un objeto “bidimensional”, luego necesitamos dos parámetros.

Llamamos **superficie paramétrica** en $\mathbb{R}^3$ a la imagen de toda función continua $\Gamma : \Omega \rightarrow \mathbb{R}^3$ donde Ω es un subconjunto no vacío, abierto y conexo de $\mathbb{R}^2$. Se trata por tanto del conjunto

$$\Sigma = \Gamma(\Omega) = \{\Gamma(t, s) : (t, s) \in \Omega\}$$

Las variables reales t y s son ahora los *parámetros*, y a cada valor $(t, s) \in \Omega$ de los mismos, corresponde un único punto $\Gamma(t, s)$ de la superficie Σ , luego podemos decir que la función Γ *parametriza* la superficie Σ . Obviamente, podemos tener otras parametrizaciones muy distintas de la misma superficie. En toda la discusión que sigue, usaremos Γ , y debe quedar claro que todo lo que hagamos depende de dicha parametrización.

Como hicimos con las curvas paramétricas, llamamos x, y, z a las tres componentes de la función Γ y decimos que

$$x = x(t, s), \quad y = y(t, s), \quad z = z(t, s), \quad ((t, s) \in \Omega)$$

son las *ecuaciones paramétricas* de la superficie $\Sigma = \Gamma(\Omega)$. Por ejemplo, el conjunto

$$\Sigma = \{ (x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1 \}$$

un cilindro vertical de radio 1, es una superficie paramétrica cuyas ecuaciones pueden ser

$$x = \cos t, \quad y = \sin t, \quad z = s, \quad (t, s \in \mathbb{R}) \quad (12)$$

Es claro que las superficies explícitas son un tipo muy particular de superficies paramétricas. Si $\Sigma = \text{Gr } f$ donde $f : \Omega \rightarrow \mathbb{R}$ es una función continua, definiendo $\Gamma(x, y) = (x, y, f(x, y))$ para todo $(x, y) \in \Omega$ tenemos una función continua $\Gamma : \Omega \rightarrow \mathbb{R}^3$ tal que $\Gamma(\Omega) = \text{Gr } f$. Digamos que una superficie explícita se puede siempre parametrizar usando como parámetros las dos primeras coordenadas de los puntos de la superficie. Podríamos usar otro par de coordenadas y considerar las superficies explícitas de la forma

$$\Sigma_1 = \{ (f(y, z), y, z) : (y, z) \in \Omega \}, \quad \text{o bien,} \quad \Sigma_2 = \{ (x, f(x, z), z) : (x, z) \in \Omega \}$$

cuyas ecuaciones explícitas serían $x = f(y, z)$ con $(y, z) \in \Omega$, o bien $y = f(x, z)$ con $(x, z) \in \Omega$.

Está claro que $\Sigma_1 = \Gamma_1(\Omega)$ y $\Sigma_2 = \Gamma_2(\Omega)$ donde

$$\Gamma_1(y, z) = (f(y, z), y, z) \quad \forall (y, z) \in \Omega \quad \text{y} \quad \Gamma_2(x, z) = (x, f(x, z), z) \quad \forall (x, z) \in \Omega$$

luego Σ_1 y Σ_2 también son superficies paramétricas. El cilindro definido en (12) es un ejemplo sencillo de superficie paramétrica que no es explícita en ninguno de los tres sentidos.

9.8. Plano tangente a una superficie paramétrica

Para extender la noción de plano tangente, fijamos una superficie paramétrica $\Sigma = \Gamma(\Omega)$ donde $\Gamma : \Omega \rightarrow \mathbb{R}^3$ es continua en un abierto conexo $\Omega \subset \mathbb{R}^2$, cuyas componentes serán x, y, z . Supongamos que Γ es diferenciable en un punto $(t_0, s_0) \in \Omega$, sea $P_0 = \Gamma(t_0, s_0) = (x_0, y_0, z_0)$ y observemos la matriz jacobiana, cuyas columnas son las derivadas parciales de Γ :

$$J\Gamma(t_0, s_0) = \begin{pmatrix} \frac{\partial x}{\partial t}(t_0, s_0) & \frac{\partial x}{\partial s}(t_0, s_0) \\ \frac{\partial y}{\partial t}(t_0, s_0) & \frac{\partial y}{\partial s}(t_0, s_0) \\ \frac{\partial z}{\partial t}(t_0, s_0) & \frac{\partial z}{\partial s}(t_0, s_0) \end{pmatrix} = \left(\frac{\partial \Gamma}{\partial t}(t_0, s_0), \frac{\partial \Gamma}{\partial s}(t_0, s_0) \right)$$

Por razones que se comprenderán enseguida, *supondremos que $J\Gamma(t_0, s_0)$ tiene rango 2*, es decir, que sus columnas son vectores linealmente independientes de $\mathbb{R}^3$. La idea clave para definir el plano tangente a Σ en el punto P_0 se puede fácilmente adivinar: las rectas tangentes en P_0 a curvas paramétricas contenidas en la superficie Σ deben estar contenidas en el plano tangente que buscamos.

Usando que Ω es abierto, podemos encontrar $\delta > 0$ de forma que se tenga $J_1 \times J_2 \subset \Omega$ donde $J_1 =]t_0 - \delta, t_0 + \delta[$ y $J_2 =]s_0 - \delta, s_0 + \delta[$. Sean entonces $\gamma_1 : J_1 \rightarrow \mathbb{R}^3$ y $\gamma_2 : J_2 \rightarrow \mathbb{R}^3$ las funciones definidas por

$$\gamma_1(t) = \Gamma(t, s_0) \quad \forall t \in J_1 \quad \text{y} \quad \gamma_2(s) = \Gamma(t_0, s) \quad \forall s \in J_2$$

Tenemos dos curvas paramétricas $C_1 = \gamma_1(J_1)$ y $C_2 = \gamma_2(J_2)$, contenidas en la superficie Σ y verificando que $P_0 = \gamma_1(t_0) = \gamma_2(s_0)$. Además γ_1 es derivable en t_0 y γ_2 en s_0 , con

$$\gamma_1'(t_0) = \frac{\partial \Gamma}{\partial t}(t_0, s_0) \quad \text{y} \quad \gamma_2'(s_0) = \frac{\partial \Gamma}{\partial s}(t_0, s_0)$$

vectores que, gracias a la hipótesis sobre la matriz $J\Gamma(t_0, s_0)$, son linealmente independientes y, en particular, no nulos. Por tanto, $P_0 = \gamma_1(t_0) = \gamma_2(s_0)$ es un punto regular de las curvas C_1 y C_2 , cuyas rectas tangentes en P_0 son distintas. Ello nos permite considerar el único plano Π que contiene a dichas rectas tangentes, es decir, el único plano que pasa por el punto P_0 y tiene como vectores de dirección a las dos derivadas parciales de Γ en (t_0, s_0) .

Por tanto las ecuaciones paramétricas del plano Π pueden escribirse en la forma:

$$\begin{aligned} x &= x_0 + (t - t_0) \frac{\partial x}{\partial t}(t_0, s_0) + (s - s_0) \frac{\partial x}{\partial s}(t_0, s_0) \\ y &= y_0 + (t - t_0) \frac{\partial y}{\partial t}(t_0, s_0) + (s - s_0) \frac{\partial y}{\partial s}(t_0, s_0) \\ z &= z_0 + (t - t_0) \frac{\partial z}{\partial t}(t_0, s_0) + (s - s_0) \frac{\partial z}{\partial s}(t_0, s_0) \end{aligned} \quad (13)$$

Nótese que obtendríamos el mismo plano escribiendo t y s en lugar de $t - t_0$ y $s - s_0$. Escribir las ecuaciones de esta forma tiene la ventaja de que el punto P_0 se obtiene tomando en ellas $t = t_0$ y $s = s_0$, igual que ocurre con las ecuaciones de la superficie Σ . Más concretamente, si para cada $(t, s) \in \Omega$ consideramos el punto $P(t, s) = (x, y, z)$ dado por estas ecuaciones, tenemos $P(t_0, s_0) = P_0$. Por supuesto, la función $P : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ parametriza el plano $\Pi = P(\Omega)$, como superficie paramétrica que es. Las ecuaciones (13) son más fáciles de recordar y de manejar si las escribimos en forma matricial:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + J\Gamma(t_0, s_0) \cdot \begin{pmatrix} t - t_0 \\ s - s_0 \end{pmatrix} \quad (14)$$

Decimos que Π es el **plano tangente** a la superficie Σ en el punto P_0 , definición que, como en el caso de una superficie explícita, se justifica de dos maneras. En primer lugar, el plano Π es una buena aproximación de la superficie Σ cerca del punto P_0 , en el sentido que sigue.

Para $(t, s) \in \Omega$, veamos la distancia del punto $\Gamma(t, s)$ de la superficie Σ al correspondiente punto $P(t, s)$ del plano Π . En vista de (14) tenemos

$$P(t, s) = \Gamma(t_0, s_0) + D\Gamma(t_0, s_0) \left((t, s) - (t_0, s_0) \right) \quad \forall t, s \in \mathbb{R}$$

y de la definición de diferenciabilidad deducimos que

$$\lim_{(t,s) \rightarrow (t_0,s_0)} \frac{\|\Gamma(t, s) - P(t, s)\|}{\|(t, s) - (t_0, s_0)\|} = 0$$

Por tanto $\|\Gamma(t, s) - P(t, s)\|$ tiende a cero cuando $(t, s) \rightarrow (t_0, s_0)$ “mucho más rápidamente” que $\|(t, s) - (t_0, s_0)\|$. Este es el tipo de buena aproximación de la superficie Σ mediante el plano Π cerca del punto P_0 que justifica analíticamente la definición de plano tangente.

Para tener una justificación más geométrica, podemos usar curvas paramétricas contenidas en la superficie Σ , más generales que las curvas C_1 y C_2 usadas para encontrar el plano Π , y comprobar que sus rectas tangentes están contenidas en Π . Concretamente, sea $\phi: J \rightarrow \Omega$ una función continua en un intervalo abierto $J \subset \mathbb{R}$, tal que $\phi(\alpha_0) = (t_0, s_0)$ para algún $\alpha_0 \in J$. La función continua $\gamma = \Gamma \circ \phi$ nos da una curva paramétrica $C = \gamma(J)$ que claramente está contenida en la superficie Σ y verifica que $\gamma(\alpha_0) = P_0$. Nótese que esta es una forma sencilla de obtener multitud de curvas paramétricas contenidas en la superficie Σ y que contienen al punto P_0 , cada función del mismo tipo que ϕ da lugar a una curva de las requeridas.

Cuando ϕ es derivable en α_0 con $\phi'(\alpha_0) \neq 0$, la regla de la cadena nos dice que γ es derivable en α_0 y, escribiendo ambos vectores derivada como matrices columna, tenemos

$$\gamma'(\alpha_0) = J\Gamma(t_0, s_0) \cdot \phi'(\alpha_0)$$

Como $J\Gamma(t_0, s_0)$ tiene rango 2, deducimos que $\gamma'(\alpha_0) \neq 0$, luego $P_0 = \gamma(\alpha_0)$ es un punto regular de la curva $C = \gamma(J)$. Escribimos las ecuaciones paramétricas de la recta tangente R , como hicimos para el plano tangente Π , es decir, las expresamos en forma matricial, haciendo que al valor $\alpha = \alpha_0$ del parámetro corresponda el punto P_0 , obteniendo

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + J\Gamma(t_0, s_0) \cdot \phi'(\alpha_0) \cdot (\alpha - \alpha_0) \quad (\alpha \in \mathbb{R}) \quad (15)$$

Está claro ahora que la recta R está contenida en el plano Π : si $(x, y, z) \in \mathbb{R}^3$ verifica (15) para algún $\alpha \in \mathbb{R}$, basta tomar $(t, s) \in \mathbb{R}^2$ dados por

$$\begin{pmatrix} t - t_0 \\ s - s_0 \end{pmatrix} = \phi'(\alpha_0) \cdot (\alpha - \alpha_0)$$

para obtener (14), luego $(x, y, z) \in \Pi$.

En resumen, para la definición de plano tangente a una superficie paramétrica, tenemos justificaciones análogas a las obtenidas para superficies explícitas: el plano tangente en cada punto P_0 , cuando existe, es una buena aproximación de la superficie cerca de P_0 , y las rectas tangentes a la superficie en dicho punto están contenidas en el plano tangente.

Resaltamos que, para definir el plano tangente a la superficie paramétrica $\Sigma = \Gamma(\Omega)$ en el punto $P_0 = \Gamma(t_0, s_0)$, y sobre todo, para poder justificarla como lo hemos hecho, es necesario suponer que Γ es diferenciable en (t_0, s_0) y que la matriz $J\Gamma(t_0, s_0)$ tiene rango 2. Cuando se cumplen estas condiciones, se dice que $P_0 = \Gamma(t_0, s_0)$ es *punto regular* de la superficie Σ , y si todos sus puntos son regulares, decimos que Σ es regular. Por tanto, $\Sigma = \Gamma(\Omega)$ es una *superficie regular* cuando Γ es diferenciable y $J\Gamma(t, s)$ tiene rango 2 para todo $(t, s) \in \Omega$. Debemos tener claro que estas nociones dependen de la parametrización Γ que usemos, para otra parametrización de la misma superficie, los puntos regulares podrían ser otros. De hecho, cuando $\Sigma = \Gamma(\Omega)$ es regular, sería más adecuado decir que Γ es una *parametrización regular* de la superficie Σ .

Nótese también que estas nociones de regularidad son análogas a las que usamos para una curva paramétrica $C = \gamma(J) \subset \mathbb{R}^M$ donde $\gamma: J \rightarrow \mathbb{R}^M$ es continua en un intervalo abierto $J \subset \mathbb{R}$. Dado $t_0 \in J$, necesitamos que γ sea derivable en t_0 con $\gamma'(t_0) \neq 0$, para tener recta tangente a la curva C en el punto $P_0 = \gamma(t_0)$. Pensemos que el vector derivada $\gamma'(t_0)$ es una matriz $M \times 1$ que tiene rango 1 cuando no se anula. Para superficies paramétricas tenemos una matriz 3×2 a la que exigimos tener rango 2.

Comprobemos que el plano tangente a una superficie explícita, considerada como superficie paramétrica, es el mismo que habíamos definido previamente. Si $f: \Omega \rightarrow \mathbb{R}$ es una función continua en un abierto $\Omega \subset \mathbb{R}^2$, la superficie explícita $\Sigma = \text{Gr } f$ se parametriza como sabemos mediante la función $\Gamma: \Omega \rightarrow \mathbb{R}^3$ dada por

$$\Gamma(x, y) = (x, y, f(x, y)) \quad \forall (x, y) \in \Omega$$

Cuando f es diferenciable en un punto $(x_0, y_0) \in \Omega$, es claro que Γ también lo es, con

$$J\Gamma(x_0, y_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}$$

luego $J\Gamma(x_0, y_0)$ tiene rango 2, y $P_0 = \Gamma(t_0, s_0) = (x_0, y_0, z_0)$ es un punto regular de Σ . Además, las ecuaciones paramétricas del plano tangente a Σ en P_0 pueden escribirse en la forma

$$x = x_0 + t, \quad y = y_0 + s, \quad z = z_0 + t \frac{\partial f}{\partial x}(x_0, y_0) + s \frac{\partial f}{\partial y}(x_0, y_0) \quad \forall t, s \in \mathbb{R}$$

luego equivalen claramente a (10). Volvemos pues a obtener el plano tangente, tal y como se definió para una superficie explícita.

Nótese que, si f es diferenciable en todo punto de Ω , entonces Γ es una parametrización regular de la superficie explícita $\Sigma = \text{Gr } f$.

Consideremos finalmente el cilindro cuyas ecuaciones paramétricas vimos en (12). En este caso, la función $\Gamma: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ es diferenciable en todo punto $(t, s) \in \mathbb{R}^2$ con

$$J\Gamma(t, s) = \begin{pmatrix} -\sin t & 0 \\ \cos t & 0 \\ 0 & 1 \end{pmatrix}$$

Claramente, $J\Gamma(t, s)$ tiene rango 2, para todo $(t, s) \in \mathbb{R}^2$, luego el cilindro $\Sigma = \Gamma(\mathbb{R}^2)$ es una superficie regular. Su plano tangente en cualquier punto $(x_0, y_0, z_0) \in \Sigma$ tiene ecuaciones paramétricas

$$x = x_0 - t y_0, \quad y = y_0 + t x_0, \quad z = z_0 + s, \quad (t, s \in \mathbb{R})$$

9.9. Curvas y superficies de nivel

Usando otra vez la regla de la cadena para las derivadas parciales, vamos a probar una propiedad del gradiente de un campo escalar que tiene interés en Física. Fijamos por tanto un campo escalar $f: U \rightarrow \mathbb{R}$, donde U es un abierto de $\mathbb{R}^N$. Llamamos conjuntos de nivel de f a los subconjuntos de U en los que toma cada uno de sus valores. Más concretamente, para cada valor $\lambda \in f(U) \subset \mathbb{R}$, el **conjunto de nivel** λ del campo f viene dado por

$$N_\lambda = \{x \in U : f(x) = \lambda\}$$

Obviamente, sin hipótesis sobre f , sus conjuntos de nivel pueden ser totalmente arbitrarios. Concentrándonos en los casos más interesantes, $N = 2$ o $N = 3$, consideramos respectivamente curvas o superficies paramétricas, que estén contenidas en un conjunto de nivel.

En el caso $N = 2$, decimos que una curva paramétrica $C = \gamma(J)$, donde $\gamma: J \rightarrow U$ es continua en un intervalo abierto $J \subset \mathbb{R}$, es una **curva de nivel** del campo f , cuando C está contenida en un conjunto de nivel de f , es decir, cuando $f \circ \gamma$ es constante.

Si γ es derivable en un punto $t_0 \in J$ y f es diferenciable en el punto $P_0 = \gamma(t_0) \in U$, la regla de la cadena nos dice que los vectores $\nabla f(P_0)$ y $\gamma'(t_0)$ son ortogonales, ya que

$$(\nabla f(P_0) | \gamma'(t_0)) = (f \circ \gamma)'(t_0) = 0$$

Claramente, esta afirmación tiene interés cuando ambos vectores son no nulos, pues entonces nos dice que *el vector gradiente $\nabla f(P_0)$ es ortogonal a la recta tangente a toda curva de nivel del campo f , para la que P_0 sea un punto regular*. Hablando con poca precisión, suele decirse que el gradiente de un campo escalar en $\mathbb{R}^2$ es ortogonal a las curvas de nivel. Si queremos desplazarnos en U de forma que el campo f se mantenga constante, es decir, siguiendo una curva de nivel, debemos hacerlo siempre en dirección ortogonal al vector gradiente.

En el caso $N = 3$, decimos que una superficie paramétrica $\Sigma = \Gamma(\Omega)$, donde $\Gamma: \Omega \rightarrow U$ es continua en un abierto conexo $\Omega \subset \mathbb{R}^2$, es una **superficie de nivel** del campo f , cuando Σ está contenida en un conjunto de nivel de f , es decir, cuando $f \circ \Gamma$ es constante. Suponiendo que Γ es diferenciable en $(t_0, s_0) \in \Omega$ y que f es diferenciable en $P_0 = \Gamma(t_0, s_0) \in U$, la regla de la cadena nos dice que el vector $\nabla f(P_0)$ es ortogonal a los vectores $\frac{\partial \Gamma}{\partial t}(t_0, s_0)$ y $\frac{\partial \Gamma}{\partial s}(t_0, s_0)$. En efecto, basta observar que

$$0 = \frac{\partial (f \circ \Gamma)}{\partial t}(t_0, s_0) = \left(\nabla f(P_0) \left| \frac{\partial \Gamma}{\partial t}(t_0, s_0) \right. \right)$$

e igual ocurre con la otra derivada parcial.

De nuevo, este resultado tiene interés cuando $\nabla f(P_0) \neq 0$ y $P_0 = \Gamma(t_0, s_0)$ es punto regular de Σ , pues entonces las dos derivadas parciales de Γ en (t_0, s_0) son vectores de dirección del plano tangente a Σ en P_0 , luego $\nabla f(P_0)$ es un vector normal a dicho plano tangente. Así pues, *el vector gradiente $\nabla f(P_0)$ es ortogonal al plano tangente a cualquier superficie de nivel del campo f , para la que P_0 sea un punto regular.* Dicho de forma imprecisa, el gradiente de un campo escalar en $\mathbb{R}^3$ es ortogonal a las superficies de nivel.

9.10. Un ejemplo ilustrativo

Hemos insistido en que todas las aplicaciones anteriores de la regla de la cadena para las derivadas parciales, sólo adquieren su verdadero sentido cuando las funciones involucradas son diferenciables. Sin embargo, cabe preguntarse si la propia regla de la cadena es válida para la composición de dos funciones parcialmente derivables, pues ello es obviamente cierto en el caso $N = M = 1$. El siguiente ejemplo, en el caso más sencillo, $N = 1$ y $M = 2$, prueba que la respuesta es negativa.

Sea $f: \mathbb{R} \rightarrow \mathbb{R}^2$ dada por $f(t) = (t, t)$ para todo $t \in \mathbb{R}$, y $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$g(x, y) = \frac{x^2 y}{x^2 + y^4} \quad \forall (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}, \quad g(0, 0) = 0$$

Obviamente f es derivable con $f'(t) = (1, 1)$ para todo $t \in \mathbb{R}$, y $g(x, 0) = g(0, y) = 0$ para cualesquiera $x, y \in \mathbb{R}$, luego g es parcialmente derivable en el origen con $\nabla g(0, 0) = (0, 0)$. Puesto que $f(0) = (0, 0)$, si la regla de la cadena para las derivadas parciales pudiera aplicarse en este caso, para la función $h = g \circ f$ se tendría $h'(0) = 0$. Vemos sin embargo que

$$h(t) = \frac{t}{1+t^2} \quad \forall t \in \mathbb{R}, \quad \text{luego} \quad h'(0) = \lim_{t \rightarrow 0} \frac{1}{1+t^2} = 1$$

Observemos que f sí es diferenciable, pero la regla de la cadena para las derivadas parciales no puede aplicarse, lo que sólo puede deberse a una razón: g no es diferenciable en el origen.

Es fácil ver que g es continua, luego $\Sigma = \text{Gr } g$ es una superficie explícita, e ingenuamente podríamos pensar que el plano de ecuación $z = 0$ es tangente a Σ en el origen. Consideremos sin embargo la curva $C = \gamma(\mathbb{R})$ donde

$$\gamma(t) = (t, t, h(t)) \quad \forall t \in \mathbb{R}$$

Tenemos claramente $C \subset \Sigma$ y el origen es un punto regular de C pero la recta tangente a C en el origen tiene vector de dirección $\gamma'(0) = (1, 1, 1)$ luego no está contenida en el supuesto plano tangente, que por tanto no merece ese nombre.

Queda claro una vez más, que la noción relevante del cálculo diferencial, desde el punto de vista teórico y por sus aplicaciones, es la diferenciabilidad, no la derivabilidad parcial.