

4 Variedades y extremos condicionados

En todo este capítulo, cada vez que hablemos de funciones \mathcal{C}^s se presupone $s \geq 1$. Vamos a definir las **variedades en \mathbb{R}^n** , las cuales, por encima de todo, tienen dos cosas: **parámetros** y **espacios tangentes**.

4.1 Difeomorfismos

Definición 101. Sean $U_1, U_2 \subseteq \mathbb{R}^n$ dos abiertos. Una aplicación $\sigma : U_1 \rightarrow U_2$ es un **difeomorfismo \mathcal{C}^s** si cumple las dos condiciones siguientes:

1. σ es biyectiva de U_1 a U_2 .
2. σ es \mathcal{C}^s y su inversa $\sigma^{-1} : U_2 \rightarrow U_1$ también es de clase \mathcal{C}^s .

Dos comentarios acerca de esta definición:

- (1) Si $\sigma : U_1 \rightarrow \mathbb{R}^n$ es *inyectiva*, de clase \mathcal{C}^s y con matrices jacobianas todas invertibles, entonces $U_2 = \sigma(U_1)$ es un abierto de \mathbb{R}^n y σ un difeomorfismo \mathcal{C}^s de U_1 a U_2 .
- (2) No toda biyección suave es un difeomorfismo. Por ejemplo $x \mapsto x^3$ es biyectiva de \mathbb{R} a \mathbb{R} y además es \mathcal{C}^∞ , pero su inversa $y \mapsto \sqrt[3]{y}$ no es diferenciable en $y = 0$ debido a la anulación de la derivada de x^3 en $x = 0$ (punto donde la jacobiana no es invertible).

Un difeomorfismo $\sigma = (\sigma_1, \dots, \sigma_n) : U_1 \rightarrow U_2$ puede entenderse como un **sistema de coordenadas curvilíneas** según explicamos a continuación. Las funciones $\sigma_1, \dots, \sigma_n : U_1 \rightarrow \mathbb{R}$ son unas **coordenadas** en U_1 en el sentido de que cada punto $p \in U_1$ está determinado por los números $\sigma_1(p), \dots, \sigma_n(p)$ y de hecho la función σ^{-1} que reconstruye el punto p a partir de esos números es de clase \mathcal{C}^s .

La definición 101 tiene perfecto sentido si suponemos, de manera más general, que U_1 es un abierto de \mathbb{R}^n y U_2 un abierto de $\mathbb{R}^{n'}$. Pero entonces, elegido un punto $a \in U_1$, la jacobiana $A = D\sigma_a$ es una matriz invertible y por lo tanto *cuadrada*. Como A es $n' \times n$ se tiene $n = n'$. Conclusión: sólo puede haber difeomorfismos entre abiertos *de igual dimensión*.

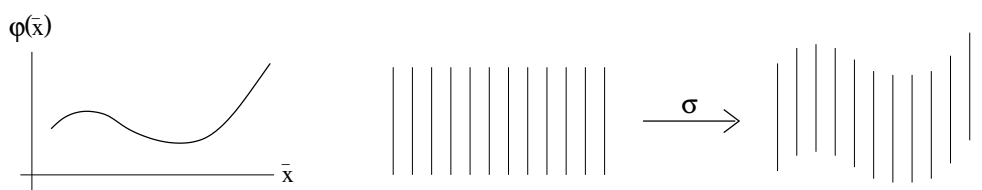
Nos interesa un tipo particular de difeomorfismos que llamaremos **deslizamientos**. Sean $k < n$, un abierto $W \subseteq \mathbb{R}^k$ y una función $\varphi = (\varphi_{k+1}, \dots, \varphi_n) : W \rightarrow \mathbb{R}^{n-k}$ con $\varphi \in \mathcal{C}^s(W)$. Para cada $x \in \mathbb{R}^n$ escribimos $\bar{x} = (x_1, \dots, x_k)$ y $\tilde{x} = (x_{k+1}, \dots, x_n)$, con lo cual $x = (\bar{x}, \tilde{x})$, y análogamente para un punto que se designe con una letra distinta de la x . Consideramos el abierto $U = \{x \in \mathbb{R}^n : \bar{x} \in W\} = W \times \mathbb{R}^{n-k}$ y la aplicación $\sigma : U \rightarrow U$ definida como sigue:

$$\sigma(\bar{x}, \tilde{x}) = (\bar{x}, \tilde{x} + \varphi(\bar{x})),$$

que es \mathcal{C}^s y biyectiva, con inversa también de clase \mathcal{C}^s :

$$\sigma^{-1}(\bar{y}, \tilde{y}) = (\bar{y}, \tilde{y} - \varphi(\bar{y})),$$

y por lo tanto σ es un difeomorfismo \mathcal{C}^s del abierto U consigo mismo. Lo que hace este difeomorfismo es deslizar cada subespacio afín $\{\bar{x}_0\} \times \mathbb{R}^{n-k} \subset \mathbb{R}^n$ dentro de sí mismo, trasladándolo por el vector $(\mathbf{0}_k, \varphi(\bar{x}_0))$.



4.2 Variedades en \mathbb{R}^n

El concepto de variedad en \mathbb{R}^n generaliza las ideas de curva y superficie, que todos entendemos de manera visual e intuitiva. De manera poco rigurosa, una curva es un objeto posiblemente curvilíneo de dimensión 1 y una superficie es un objeto posiblemente curvilíneo de dimensión 2. La generalización que utilizaremos aquí permite otras dimensiones (no sólo 1 o 2) y exige una cierta “buena calidad” a los objetos definidos.

Definición 102. Sean k un entero no negativo y $X \subseteq \mathbb{R}^n$ un subconjunto no vacío. Decimos que X es una **variedad en \mathbb{R}^n de dimensión geométrica k** si para cada punto $x^0 \in X$ existen abiertos $E, E' \subseteq \mathbb{R}^n$, con $x^0 \in E$, y un difeomorfismo $\sigma : E \rightarrow E'$ tales que:

$$\sigma(X \cap E) = E' \cap (\mathbb{R}^k \times \{\mathbf{0}_{n-k}\}) = \{y \in E' : y_{k+1} = y_{k+2} = \dots = x_n = 0\}. \quad (47)$$

Veamos algunos valores especiales de la dimensión geométrica k . Cuando $k = 1$ decimos que X es una **curva en \mathbb{R}^n** . Cuando $k = 2$ decimos que X es una **superficie en \mathbb{R}^n** . Las variedades en \mathbb{R}^n de dimensión $n - 1$ se llaman **hipersuperficies**. Las variedades de dimensión 0 son los conjuntos **discretos**: los que sólo tienen puntos aislados. Las variedades de dimensión n en \mathbb{R}^n son los abiertos no vacíos.

Esta definición encierra mucho más de lo que parece, por lo cual tenemos que hacerle varios comentarios.

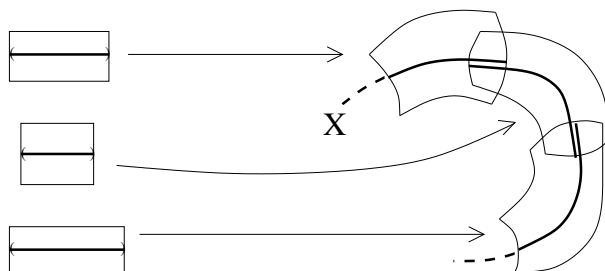
1. La palabra “en” forma parte del nombre de estos objetos, incluso cuando ella y la palabra “variedad” no vayan juntas, como por ejemplo al decir “**variedad de dimensión 2 en \mathbb{R}^3** ”. Hemos definido, pues, el concepto de “variedades en”. En un curso más avanzado se definirá un concepto más general de variedades que no necesariamente están en ningún \mathbb{R}^n , y esas ya no serán “variedades en”, se las llamará simplemente “variedades”.

2. Podemos considerar a $\mathbb{R}^k \times \{\mathbf{0}_{n-k}\}$ como la manera estándar de meter \mathbb{R}^k dentro de \mathbb{R}^n como un subespacio afín. Además es $E' \cap (\mathbb{R}^k \times \{\mathbf{0}_{n-k}\}) = V \times \{\mathbf{0}_{n-k}\}$, donde V es un abierto de \mathbb{R}^k , luego podemos considerar a $E' \cap (\mathbb{R}^k \times \{\mathbf{0}_{n-k}\})$ como una manera estándar de meter (sin curvarlo) un cierto abierto V de \mathbb{R}^k dentro del espacio ambiente \mathbb{R}^n .

3. Es importante que en (47) tengamos igualdad de conjuntos: si la reemplazamos por la inclusión \subseteq entonces se permite que la imagen $\sigma(X \cap E)$ pueda ser $C \times \{\mathbf{0}_{n-k}\}$ donde $C \subset \mathbb{R}^k$ no sea un abierto de \mathbb{R}^k sino un subconjunto muy extraño. Tal situación la hemos prohibido con la igualdad conjuntista en (47):

Estamos exigiendo que partes pequeñas de X sean como una bola abierta de \mathbb{R}^k

4. Mientras que $\sigma : E \rightarrow E'$ es un *difeomorfismo planchador*, porque toma el trocito $X \cap E$ y lo plancha (lo convierte en un abierto del subespacio $\mathbb{R}^k \times \{\mathbf{0}_{n-k}\}$), el inverso $\sigma^{-1} : E' \rightarrow E$ es un *difeomorfismo curvador*: toma el trozo abierto $E' \cap (\mathbb{R}^k \times \{\mathbf{0}_{n-k}\})$ de un subespacio afín y lo curva de modo a construir el “trozo curvado” $X \cap E$. El conjunto total X está formado por estos trocitos curvados que, al ser abiertos relativos de X , no van a poder ser disjuntos dos a dos en cuanto X sea conexo por caminos. Ahora bien, cuando dos abiertos de un espacio métrico comparten un punto a automáticamente comparten toda una bola $B(a, \varepsilon)$, es decir que *se solapan*, luego los trocitos curvados $X \cap E$ van a tener solapamientos entre ellos.



Definición 103. Decimos que X es de clase \mathcal{C}^s , como variedad en \mathbb{R}^n , si para ella podemos elegir, cerca de cada punto, difeomorfismos planchadores de clase \mathcal{C}^s .

Ejemplos de variedades en \mathbb{R}^n :

(1) Subespacios afines, o abiertos relativos suyos. Se los plancha con **movimientos** de \mathbb{R}^n . Dado un subespacio afín k -dimensional $\mathbb{A} \subset \mathbb{R}^n$ y dado un abierto relativo $V \subseteq \mathbb{A}$, tenemos una igualdad $V = \mathbb{A} \cap U$ donde U es un abierto de \mathbb{R}^n . Existe un movimiento $M(x) = b + Ax$, con $b \in \mathbb{R}^n$ vector constante y A matriz ortogonal constante, tal que $M(\mathbb{A}) = \mathbb{R}^k \times \{\mathbf{0}_{n-k}\}$. Entonces $\sigma = M|_U$ es un difeomorfismo entre abiertos de \mathbb{R}^n que mueve V y lo transforma en $V' \times \{\mathbf{0}_{n-k}\}$ para algún abierto $V' \subseteq \mathbb{R}^k$.

(2) Grafos de funciones diferenciables. Se los plancha con deslizamientos. Hay **varios tipos de grafo**: se eligen $n - k$ variables y se las pone como funciones \mathcal{C}^s de las otras k variables, dejando que éstas últimas recorran un abierto de \mathbb{R}^k . Por ejemplo, dado un abierto $W \subseteq \mathbb{R}_{x_1 x_3}^2$ y un grafo en \mathbb{R}^4 del siguiente tipo:

$$X = \{ (x_1, \varphi_2(x_1, x_3), x_3, \varphi_4(x_1, x_3)) : (x_1, x_3) \in W \},$$

en el cual x_2, x_4 están puestas como funciones diferenciables de (x_1, x_3) , definimos el abierto $U = \{x \in \mathbb{R}^4 : (x_1, x_3) \in W\}$ y X se plancha con el deslizamiento $\sigma : U \rightarrow U$ dado por:

$$(x_1, x_2, x_3, x_4) \xrightarrow{\sigma} (x_1, x_2 - \varphi(x_1, x_3), x_3, x_4 - \varphi_4(x_1, x_3)).$$

Veamos ahora ejemplos de subconjuntos que no son variedades en \mathbb{R}^n .

El conjunto $T = \{(x, y) : xy = 0, y \geq 0\}$, formado por tres semirrectas que salen del origen, no es una variedad en el plano.

Los grafos de $|x|$ y de $2x + |x|$ no son variedades en el plano. Los conos $\{y = \sqrt{x^2 + y^2}\}$ y $\{x^2 + y^2 - z^2 = 0\}$ no son variedades en \mathbb{R}^3 . Véanse los comentarios después de la proposición 105, en el apartado 4.3.

4.3 Funciones regulares: preimágenes e imágenes

Sean $U \subseteq \mathbb{R}^n$ un abierto y $f : U \rightarrow \mathbb{R}^m$ una función al menos de clase \mathcal{C}^1 . En la definición 87 del apartado 3.6 hemos dicho que f es **regular** si para todo $a \in U$ la diferencial $(df)_a$ tiene el máximo rango que puede tener una función lineal $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Lo separamos en tres posibilidades:

1. Si $n < m$ entonces el máximo rango posible para L es n y lo tienen las L inyectivas. En este caso, pues, la función f es regular si todas sus diferenciales son inyectivas. Esto también se expresa diciendo que f es una **inmersión** de U dentro de \mathbb{R}^m . La matriz Df tiene menos columnas que filas y la condición para que f sea una inmersión es que las columnas f_{x_1}, \dots, f_{x_n} de Df sean linealmente independientes en todo punto. En particular, un **camino** $\alpha(t)$ es una inmersión si el vector velocidad $\alpha'(t)$ no se anula para ningún t .
2. Si $n = m$ volvemos al caso estudiado en el capítulo 3. Ahora f es regular si sus diferenciales son biyectivas; las matrices Df son cuadradas e invertibles. Cada punto $a \in U$ tiene un entorno U^a tal que $f : U^a \rightarrow f(U^a)$ es un difeomorfismo. No obstante f puede no ser inyectiva en todo U . A una tal f también se la llama **difeomorfismo local**.
3. Si $n > m$ entonces el máximo rango posible para L es m y lo tienen las L suprayectivas. En este caso, pues, la función f es regular si todas sus diferenciales son suprayectivas. A una tal f también se la llama **submersión**. La matriz Df tiene menos filas que columnas y alcanza rango m si dichas filas son linealmente independientes. Escribiendo

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix}, \text{ dichas filas son } Df_1, \dots, Df_m, \text{ equiparables a los } \mathbf{gradientes} \nabla f_1, \dots, \nabla f_m.$$

La función f es una submersión si esos gradientes son linealmente independientes en todo punto.

Teorema 104. (Preimágenes). Sean $U_0 \subseteq \mathbb{R}^n$ un abierto, $n \geq m$ y $f : U_0 \rightarrow \mathbb{R}^m$ de clase \mathcal{C}^s .

Dado $b \in \mathbb{R}^m$, si $(df)_a$ es suprayectiva para todo $a \in f^{-1}(\{b\})$, entonces $f^{-1}(\{b\})$ es o el conjunto vacío o una variedad en \mathbb{R}^n de dimensión $n - m$ y clase \mathcal{C}^s .

Veamos otra manera de expresar este teorema. Dado un sistema

$$\left. \begin{array}{l} f_1(x_1, \dots, x_n) = b_1 \\ \dots \\ f_m(x_1, \dots, x_n) = b_m \end{array} \right\} (x_1, \dots, x_n) \in U_0,$$

si los gradientes de las ecuaciones $\nabla f_1, \dots, \nabla f_m$ son linealmente independientes en cada punto solución del sistema, entonces el conjunto X de las soluciones es o vacío o una variedad en \mathbb{R}^n y además $\dim X = n - 1$ si hay una ecuación, $\dim X = n - 2$ si hay dos ecuaciones, etc.

Una sola ecuación escalar $f_1(x_1, \dots, x_n) = b_1$, con al menos una solución y con gradiente no nulo en todo punto solución, define una hipersuperficie en \mathbb{R}^n . Por ejemplo el conjunto

$$S^{n-1} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : x_1^2 + \dots + x_n^2 = 1\},$$

es una variedad en \mathbb{R}^n de dimensión $n - 1$ y se la llama **esfera $(n - 1)$ -dimensional**. El gradiente $\nabla(x_1^2 + \dots + x_n^2) = (2x_1, \dots, 2x_n)$ se anula en el origen $\mathbf{0}$, pero éste es un punto en el que la ecuación no se satisface y por lo tanto no causa ningún problema. En particular, la esfera S^2 es una superficie en \mathbb{R}^3 y la circunferencia S^1 es una curva en \mathbb{R}^2 .

Demostración del teorema 104. Sea $X = f^{-1}(\{b\}) = \{x \in U_0 : f(x) = b\}$. Supongamos X no vacío y fijemos un punto $a \in X$.

Primer caso: $n = m$. La matriz Df_a es invertible y existen entornos U de a y V de b tales que f es biyectiva de U a V . Se sigue que

$$U \cap X = \{x \in U : f(x) = b\} = \{a\}.$$

luego a es un punto aislado de X . Como esto es así para todo punto suyo, el conjunto X es discreto, o sea una variedad en \mathbb{R}^n de dimensión $0 = n - n$.

Segundo caso: $n > m$. Las m filas de la jacobiana Df_a son linealmente independientes y sabemos, por Álgebra Lineal, que hay una submatriz invertible $A_{m \times m}$ de Df_a . Consideramos la correspondiente partición

$$\{1, \dots, n\} = \{i_1, \dots, i_{n-m}\} \sqcup \{j_1, \dots, j_m\},$$

en la que j_1, \dots, j_m son los índices de las columnas de A e i_1, \dots, i_{n-m} los índices restantes. Escribimos $k = n - m$ y pasamos a considerar las variables x_{i_1}, \dots, x_{i_k} como parámetros, definiendo una familia paramétrica de funciones de m variables:

$$F_{(x_{i_1}, \dots, x_{i_k})}(x_{j_1}, \dots, x_{j_m}) = f(x_1, \dots, x_n).$$

Por ejemplo, si $n = 5$ y si la matriz A está formada por las columnas *primera* y *tercera* de Df entonces reinterpretamos la función original de cinco variables como una familia a tres parámetros de funciones de dos variables (y con valores en \mathbb{R}^2) de la siguiente manera:

$$F_{(x_2, x_4, x_5)}(x_1, x_3) = f(x_1, x_2, x_3, x_4, x_5).$$

A cada $x \in \mathbb{R}^n$ le asociamos los vectores:

$$\bar{x} = (x_{i_1}, \dots, x_{i_k}) \quad , \quad \tilde{x} = (x_{j_1}, \dots, x_{j_m}).$$

y análogamente si el vector se designa por una letra distinta de la x . El vector (\bar{x}, \tilde{x}) es el resultado de hacer una permutación en las entradas de (x_1, \dots, x_n) .

La matriz $(DF_{\bar{a}})_{\bar{a}} = (D_{x_{j_1}, \dots, x_{j_m}} f)_a$ es la submatriz A que hemos supuesto invertible, por lo tanto a la familia $F_{\bar{x}}$ se le aplica el teorema 98 de las funciones implícitas: existen entornos W de \bar{a} en \mathbb{R}^k y U de \tilde{a} en \mathbb{R}^m , tales que si definimos un abierto $E \subseteq \mathbb{R}^n$ como sigue:

$$x \in E \iff (\bar{x} \in W \text{ y } \tilde{x} \in U) \iff (\bar{x}, \tilde{x}) \in W \times U,$$

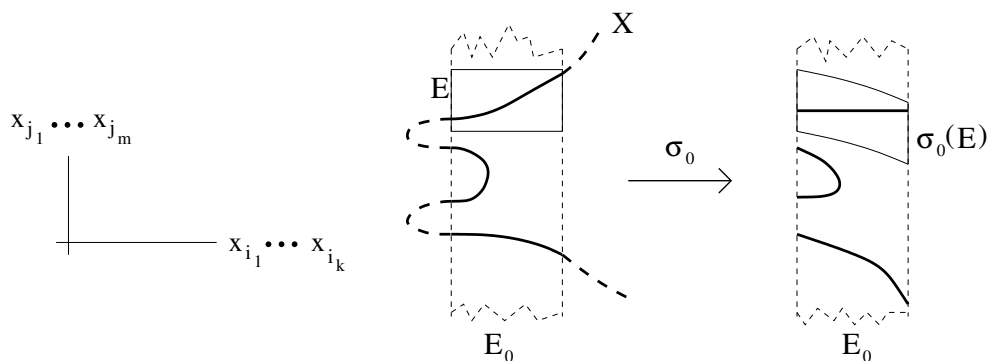
entonces $E \subseteq U_0$ y la parte del conjunto X contenida en E tiene la siguiente descripción:

$$x \in X \cap E \iff (\bar{x} \in W \text{ y } \tilde{x} = \varphi(\bar{x})),$$

para cierta función $\varphi : W \rightarrow U$ que es de la misma clase \mathcal{C}^s que la función f . Esto significa que $X \cap E$ es un **grafo** en el que las variables $(x_{j_1}, \dots, x_{j_m})$ son puestas como funciones \mathcal{C}^s de los parámetros $(x_{i_1}, \dots, x_{i_k})$. Podemos definir la aplicación $\sigma : E \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^s dada por:

$$\bar{\sigma}(x) = \bar{x} \quad , \quad \tilde{\sigma}(x) = \tilde{x} - \varphi(\bar{x}).$$

Atención. En realidad σ es la restricción a E de un deslizamiento σ_0 que está definido en el abierto más grande $E_0 \subseteq \mathbb{R}^n$ definido por $E_0 = \{x \in \mathbb{R}^n : \bar{x} \in W\} \supseteq E$. Es importante restringir σ_0 al abierto más pequeño E , porque el conjunto total X puede cortar a E_0 en otras partes distintas de $X \cap E$, y esas tal vez no las planche σ_0 .



El conjunto $\sigma(E)$, imagen directa de E por σ_0 , es un abierto de \mathbb{R}^n y tenemos un difeomorfismo $\sigma : E \rightarrow \sigma(E)$ de clase \mathcal{C}^s y tal que:

$$\sigma(X \cap E) = \{y \in \mathbb{R}^n : \bar{y} \in W, \tilde{y} = \mathbf{0}_m\} = \{y \in \sigma(E) : \tilde{y} = \mathbf{0}_m\},$$

es decir que σ es un difeomorfismo que plancha la parte $X \cap E$ de X : la convierte en un abierto del subespacio afín $\{y : \tilde{y} = \mathbf{0}_m\}$, que tiene dimensión $k = n - m$. Al ser esto posible para todo $a \in X$, concluimos que X es una variedad en \mathbb{R}^n de dimensión geométrica $n - m$ y clase \mathcal{C}^s . \square

En la demostración del teorema 104 hemos considerado varios tipos de grafo: tantos como subconjuntos $\{i_1, \dots, i_k\}$ hay en $\{1, \dots, n\}$. Veamos un ejemplo en \mathbb{R}^4 en el que aparecen grafos de distinto tipo. Consideramos la función $f : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ definida como sigue:

$$f(x_1, x_2, x_3, x_4) = \begin{bmatrix} x_2 \cos x_1 + x_3^2 + 7x_2x_4 \\ e^{x_1}x_3 + 5e^{x_2} - \text{sen } x_3 - x_1x_4^2 \end{bmatrix},$$

que es claramente \mathcal{C}^∞ . Dados los puntos $a = (0, 1, 0, 0)$, $a' = (0, 0, 0, 1)$, $a'' = (0, 1, 1, 0)$ y sus respectivas imágenes:

$$b = f(a) = \begin{bmatrix} 1 \\ 5e \end{bmatrix}, \quad b' = f(a') = \begin{bmatrix} 0 \\ 5 \end{bmatrix}, \quad b'' = f(a'') = \begin{bmatrix} 2 \\ 1 + 5e - \text{sen } 1 \end{bmatrix},$$

queremos estudiar el conjunto $f^{-1}(\{b\})$ en las cercanías de a , el conjunto $f^{-1}(\{b'\})$ en las cercanías de a' y el conjunto $f^{-1}(\{b''\})$ en las cercanías de a'' .

Empezamos por el punto a . La preimagen $f^{-1}(\{b\})$ es el conjunto de soluciones $x = (x_1, x_2, x_3, x_4)$ al sistema $f(x) = b$, es decir:

$$\left. \begin{aligned} x_2 \cos x_1 + x_3^2 + 7x_2x_4 &= 1 \\ e^{x_1}x_3 + 5e^{x_2} - \sin x_3 - x_1x_4^2 &= 5e \end{aligned} \right\}, \quad (48)$$

y de entre ellas queremos describir *solamente* las x cercanas al punto a . Calculamos:

$$Df = \begin{bmatrix} -x_2 \sin x_1 & \cos x_1 + 7x_4 & 2x_3 & 7x_2 \\ e^{x_1}x_3 - x_4^2 & 5e^{x_2} & e^{x_1} - \cos x_3 & -2x_1x_4 \end{bmatrix}.$$

En particular:

$$Df_a = \begin{bmatrix} 0 & 1 & 0 & 7 \\ 0 & 5e & 0 & 0 \end{bmatrix}_{2 \times 4}.$$

Esta matriz tiene $\binom{4}{2} = 6$ submatrices 2×2 , de las cuales solamente es invertible la formada por las columnas segunda y cuarta. Tenemos garantizado, por el teorema de las funciones implícitas, que las soluciones de (48) *cercanas al punto a* vienen dadas por

$$(x_2, x_4) = (\varphi_2(x_1, x_3), \varphi_4(x_1, x_3)),$$

donde $\varphi = (\varphi_2, \varphi_4)$ es una función vectorial C^∞ y (x_1, x_3) recorre un entorno de $(a_1, a_3) = (0, 0)$ en el plano $\mathbb{R}_{x_1x_3}^2$. El teorema de las funciones implícitas *no dice* si cerca del punto a se puede o no se puede despejar en (48) una pareja de variables distinta de la (x_2, x_4) como función de las otras dos variables; decidirlo requiere un estudio de las ecuaciones (48) que va más allá de las derivadas primeras de f en a y que no vamos a hacer.

Para el punto $a' = (0, 0, 0, 1)$ calculamos:

$$Df_{a'} = \begin{bmatrix} 0 & 8 & 0 & 0 \\ -1 & 5 & 0 & 0 \end{bmatrix}.$$

La única submatriz invertible 2×2 es la formada por las dos primeras columnas. El teorema de las funciones implícitas dice que cerca de a' se pueden despejar en el sistema $f(x) = b'$ las variables (x_1, x_2) como funciones de clase C^∞ de (x_3, x_4) que a su vez recorren un entorno de $(a'_3, a'_4) = (0, 1)$ en el plano $\mathbb{R}_{x_3x_4}^2$.

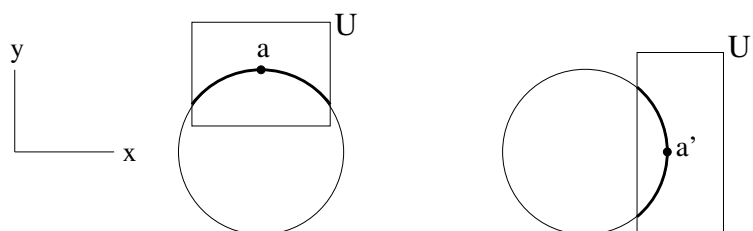
El teorema de las funciones implícitas no garantiza (tampoco prohíbe) que cerca del punto a' podamos despejar en $f(x) = b'$ una pareja de variables distinta de la x_1, x_2 como función C^∞ de las otras dos variables.

Para el punto $a'' = (0, 1, 1, 0)$ calculamos:

$$Df_{a''} = \begin{bmatrix} 0 & 1 & 2 & 7 \\ 1 & 5e & 1 - \cos 1 & 0 \end{bmatrix}.$$

ahora las seis submatrices 2×2 son invertibles. El teorema de las funciones implícitas garantiza que cerca de a'' podemos despejar en $f(x) = b''$ cualquier pareja de variables (x_i, x_j) como función C^∞ de las otras dos (x_k, x_l) que a su vez recorren un entorno de (a''_k, a''_l) en el plano $\mathbb{R}_{x_kx_l}^2$.

Una sola variedad, incluso conexa por caminos, puede obligarnos a cambiar el tipo de grafo de un punto a otro. Para la circunferencia $X = \{(x, y) : x^2 + y^2 = 1\}$ y entornos U del punto $a = (0, 1)$, la parte $X \cap U$ sólo puede describirse poniendo y como función de x y no x como función de y . En cambio para entornos U' del punto $a' = (1, 0)$ la parte $X \cap U'$ sólo puede describirse poniendo x como función de y .



Proposición 105. Sea X una variedad en \mathbb{R}^n de dimensión k y clase \mathcal{C}^s . Cada punto $a \in X$ tiene un entorno E tal que la parte $X \cap E$ es un grafo donde $n - k$ de las variables x_1, \dots, x_n son puestas como funciones de clase \mathcal{C}^s de las otras k variables y éstas últimas recorren un abierto de \mathbb{R}^k .

Demostración. Sea $a \in X$. Tenemos un entorno E_1 de a y un difeomorfismo $\sigma : E_1 \rightarrow E_2$ tales que $\sigma(X \cap E_1) = \{y \in E_2 : y_{k+1} = \dots = y_n = 0\}$. Por lo tanto:

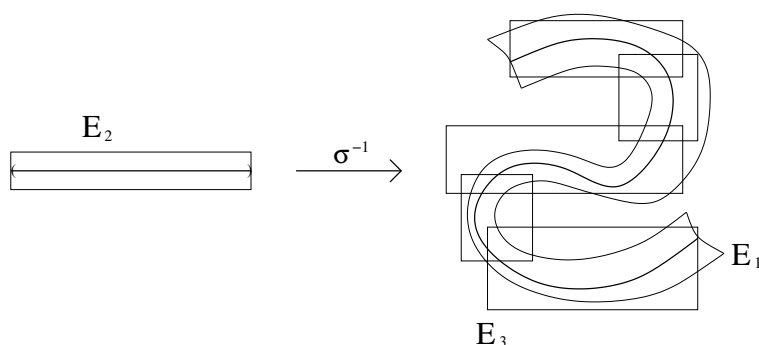
$$X \cap E_1 = \{x \in E_1 : \sigma_{k+1}(x) = \dots = \sigma_n(x) = 0\}.$$

Como σ es un difeomorfismo entre abiertos de \mathbb{R}^n , sus jacobianas son invertibles y tienen todas las filas linealmente independientes. Al ser $D\sigma_{k+1}, \dots, D\sigma_n$ algunas de las filas de $D\sigma$, resulta que son linealmente independientes en todo punto de E_1 . El sistema

$$\left. \begin{array}{l} \sigma_{k+1}(x) = 0 \\ \vdots \\ \sigma_n(x) = 0 \end{array} \right\} x \in E_1,$$

que define la parte $X \cap E_1$, está en las hipótesis del teorema 104 y, por la demostración que hemos hecho del mismo, para cada punto $a \in X \cap E_1$ existe un entorno E_3 tal que la parte $X \cap E_1 \cap E_3$ es un grafo del algún tipo con k variables independientes. \square

Dada una variedad X en \mathbb{R}^n , y una parte $X \cap E_1$ que se planche por un difeomorfismo $\sigma : E_1 \rightarrow E_2$, pueden necesitarse grafos de distinto tipo para recubrir $X \cap E_1$, siendo inevitable, para recubrirla así, descomponerla en muchos trocitos $X \cap E_1 \cap E_3$.



Veamos que el grafo de $X = \{y = |x|\}$ no es una variedad en el plano. Su dimensión no puede ser 0 porque no es un conjunto discreto, tampoco puede ser 2 porque no es un abierto. La proposición 105 afirma que, si fuera una variedad unidimensional en el plano, entonces la parte cercana al punto $(0, 0) \in X$ sería un grafo en el que o bien y sería función diferenciable de x o bien x sería función diferenciable de y . Es evidente que en $X \cap B((0, 0), \varepsilon)$ sólo puede ponerse y como función de x y que sale una función no diferenciable, luego X no es una variedad en el plano.

En particular, no existe ningún difeomorfismo entre abiertos del plano que lleve X al eje de abscisas. Intuitivamente, esto corresponde al hecho de que X “tiene una esquina” en el punto $(0, 0)$. La definición que hemos dado del concepto de variedad en \mathbb{R}^n es adecuada porque **detecta la presencia de esquinas**; era delicado conseguir esto, porque existen biyecciones diferenciables entre abiertos del plano que planchan X , por ejemplo f dada por $f(x, y) = (x, y^5 - |x|^5)$, que es \mathcal{C}^4 y, aunque biyectiva de \mathbb{R}^2 a \mathbb{R}^2 , tiene $Df_{(0,0)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ no invertible.

Por el mismo argumento, el grafo $\{z = \sqrt{x^2 + y^2}\}$ (cono simple) no es una variedad en \mathbb{R}^3 . Tampoco lo es el cono doble $\{x^2 + y^2 - z^2 = 0\}$, porque cerca del origen (que es un punto suyo) no es un grafo de ninguna de las maneras que se contemplan en la proposición 105.

El grafo de $Y = \{y = 2x + |x|\}$ no es una variedad en el plano. De hecho sí es un grafo de las dos maneras (y función de x y x función de y) pero resultan dos funciones no diferenciables.

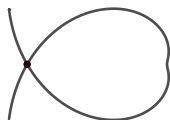
Teorema 106. (Imágenes). Sea $k < n$. Sean $V \subseteq \mathbb{R}^k$ un abierto y $f : V \rightarrow \mathbb{R}^n$ una **inmersión** de clase C^s , es decir una aplicación de clase C^s con las diferenciales todas invertidas. La imagen $f(V)$ es una unión de variedades de clase C^s y dimensión k en \mathbb{R}^n .

Demostración. Fijamos $x^0 \in V$. La matriz Df_{x^0} tiene rango k y por lo tanto k filas linealmente independientes. Sean i_1, \dots, i_k los índices de dichas filas y j_1, \dots, j_m los demás índices, siendo $m = n - k$. Entonces $\sigma = (f_{i_1}, \dots, f_{i_k})$ tiene jacobiana invertible en x^0 y es un difeomorfismo de un entorno $V^{x^0} \subseteq V$ a un entorno W de $\bar{y}_0 = \sigma(x^0)$ en \mathbb{R}^k . La imagen $f(V^{x^0})$ es el siguiente grafo:

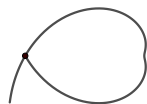
$$\{ y : \bar{y} = (y_{i_1}, \dots, y_{i_k}) \in W, (y_{j_1}, \dots, y_{j_m}) = (f_{j_1} \circ \sigma^{-1}(\bar{y}), \dots, f_{j_m} \circ \sigma^{-1}(\bar{y})) \}.$$

Esto define un recubrimiento de la imagen total $f(V)$ por imágenes $f(V^{x^0})$, que son grafos de funciones de clase C^s definidas en abiertos de \mathbb{R}^k . Tales grafos son variedades de clase C^s y dimensión geométrica k . \square

Aviso. Es bastante delicado decidir si la imagen de una inmersión es una sola variedad en \mathbb{R}^n o unión de varias de ellas; veamos unos ejemplos. El camino $\alpha(t) = ((1+t^2)\cos t, (1+t^2)\sin t)$ no es inyectivo en $t \in (-3'5, 3'5)$ y la imagen $\alpha((-3'5, 3'5))$ tiene un “cruce consigo misma” en el punto $p = \alpha(-\pi) = \alpha(\pi)$.



Esta imagen es unión de dos variedades en el plano, no una sola. A pesar de que α es inyectivo en $t \in (-\pi, 3'5)$, la imagen más pequeña $\alpha((-\pi, 3'5))$ también es unión de por lo menos dos variedades en el plano, porque en el punto p presenta el mismo problema que el ejemplo T del final del apartado 4.2.



El camino $\beta(t) = (\cos t, \sin t)$, $-\infty < t < \infty$, es una inmersión no inyectiva y, sin embargo, su imagen $\beta(\mathbb{R})$ sí es una variedad en el plano.

4.4 Dimensión y grado de suavidad

La dimensión geométrica de una variedad X en \mathbb{R}^n es única. Si X tiene dos dimensiones k, k' cerca de un punto $a \in X$, entonces se construye un difeomorfismo $U \rightarrow U'$ entre abiertos de \mathbb{R}^n que lleva $U \cap (\mathbb{R}^k \times \{\mathbf{0}_{n-k}\})$ a $U' \cap (\mathbb{R}^{k'} \times \{\mathbf{0}_{n-k'}\})$. Así se induce un difeomorfismo de un abierto de \mathbb{R}^k a un abierto de $\mathbb{R}^{k'}$, lo cual fuerza $k = k'$. Esto prueba que la dimensión es única en un punto dado, pero no quita para que pueda ser distinta en otro punto como ocurre, por ejemplo, en el conjunto $\mathcal{X} = \{y = z = 0\} \cup \{z = 1\} \subset \mathbb{R}^3$ que es la unión disjunta de una recta y un plano en \mathbb{R}^3 . Pues bien: \mathcal{X} no es una variedad en \mathbb{R}^3 por definición. Si miramos despacio la definición 102, dice que hay fijado de antemano un entero k que afecta a todos los puntos del conjunto en cuestión. Podemos llamar a \mathcal{X} “unión de dos variedades en \mathbb{R}^3 ”, pero no llamarlo “variedad en \mathbb{R}^3 ” tal como está establecida la definición.

Proposición 107. Sea $W \subseteq \mathbb{R}^k$ un abierto. Si $\varphi : W \rightarrow \mathbb{R}^m$ es de clase C^1 y su grafo es una variedad de clase C^s en \mathbb{R}^{k+m} , entonces φ es una función de clase C^s .

Importante. Se necesita, para poder aplicar la proposición, que φ sea al menos C^1 . Por ejemplo, el grafo de la función $y = \sqrt[3]{x}$ es una variedad de clase C^∞ en el plano, pero no se puede aplicar la proposición y concluir que la función sea C^∞ (no lo es) porque ni siquiera es C^1 .

Demostración parcial. Lo probamos sólo en el caso $k = m = 1$, en el cual el grafo es una variedad unidimensional en \mathbb{R}^2 . Por la demostración de la proposición 105, dado un punto $a = (x^0, y^0)$ de dicho grafo hay un entorno $E_1 \subseteq \mathbb{R}^2$ de a y una función escalar $f \in \mathcal{C}^s(E_1)$ tal que $\nabla f_a \neq (0, 0)$ y la parte del grafo contenida en E_1 puede darse como $\{(x, y) \in E_1 : f(x, y) = 0\}$. Entonces tenemos la identidad $f(x, \varphi(x)) = 0$ en un abierto de \mathbb{R}_x que contiene a x^0 . Como estamos suponiendo $\varphi \in \mathcal{C}^1$, se puede derivar esa identidad en $x = x^0$ y obtener la siguiente igualdad:

$$f_x(x^0, y^0) + \varphi'(x^0) f_y(x^0, y^0) = 0,$$

es decir $f_x(a) + \varphi'(x^0) f_y(a) = 0$. Se deduce que si fuera $f_y(a) = 0$ entonces también sería $f_x(a) = 0$ y tendríamos $\nabla f_a = (0, 0)$, lo cual es falso. Luego $f_y(a) \neq 0$ y, por el teorema de las funciones implícitas y la identidad $f(x, \varphi(x)) = 0$, la función $\varphi(x)$ es de clase \mathcal{C}^s cerca de x^0 . Al ser esto así para todo punto (x^0, y^0) del grafo, la función $\varphi(x)$ es de clase \mathcal{C}^s en todo su dominio. \square

La proposición 107 dice, entre otras cosas, que la definición 103 no está vacía de contenido: para cada $s \geq 1$ finito hay variedades en \mathbb{R}^n que son de clase \mathcal{C}^s y no de clase \mathcal{C}^{s+1} ; basta tomar el grafo de cualquier función que sea \mathcal{C}^s y no sea \mathcal{C}^{s+1} .

Por ejemplo, el grafo $\{y = |x|^{3/2}\}$ lo podemos planchar con un difeomorfismo del plano (un deslizamiento) que es de clase \mathcal{C}^1 , pero es imposible hacerlo con difeomorfismos entre abiertos del plano que sean de clase \mathcal{C}^2 o mejor.

4.5 Parámetros

A continuación definimos unos objetos que utilizamos para casi todo cuando manejamos una variedad.

Definición 108. Sea X una variedad en \mathbb{R}^n de dimensión k . Una **parametrización** para X viene dada por un abierto $W \subseteq \mathbb{R}^k$ y una función $\Phi : W \rightarrow \mathbb{R}^n$ cumpliendo las siguientes condiciones:

1. Φ es de clase \mathcal{C}^s .
2. Φ toma valores sólo en X , es decir $\Phi(W) \subseteq X$.
3. La imagen $\Phi(W)$ tiene interior no vacío en X : contiene un abierto relativo $\emptyset \neq Y \subseteq X$.

Si (u_1, \dots, u_k) , denota el punto general de W , entonces las variables u_1, \dots, u_k son los **parámetros**. Llamamos **parte parametrizada por Φ** a la imagen $\Phi(W)$.

Dos comentarios:

(1) No le exigimos a una parametrización que recorra todo el conjunto X , pero la idea detrás de la condición 3. es que $\Phi(W)$ no sea una “parte insignificante” de X . Por ejemplo, si X es una superficie en \mathbb{R}^3 entonces pedimos que $\Phi(W)$ contenga una parte de X que también sea una superficie.

(2) Algunas parametrizaciones tienen mayor calidad que otras, siendo las mejores las que son a la vez inyectivas y regulares. En este caso “regular” quiere decir que la parametrización es una inmersión, o sea con todas las diferenciales inyectivas.

Veamos unos ejemplos. La función $\Phi : \mathbb{R}_{\theta, \varphi}^2 \rightarrow \mathbb{R}_{xyz}^3$ dada por:

$$\Phi(\theta, \varphi) = (\cos \varphi \cos \theta, \cos \varphi \sin \theta, \sin \varphi) \quad , \quad \Phi \in \mathcal{C}^\infty(\mathbb{R}^2) \quad , \quad (49)$$

es una “fábrica de soluciones” de la ecuación $x^2 + y^2 + z^2 = 1$: para cualesquiera números θ, φ , el punto $(x, y, z) = \Phi(\theta, \varphi)$ satisface esa ecuación. De este modo Φ nos permite movernos por la esfera S^2 : a medida que cambia el valor paramétrico (θ, φ) , el punto $\Phi(\theta, \varphi)$ se mueve por la esfera S^2 sin salirse de ella. Llegamos a cualquier punto de la esfera porque $\Phi(\mathbb{R}^2) = S^2$.

La parametrización Φ dista mucho de ser inyectiva: cada punto $p \in S^2$ tiene infinitas preimágenes por Φ . Tampoco es regular, ya que su jacobiana:

$$D\Phi_{(\theta,\varphi)} = [\Phi_\theta | \Phi_\varphi] = \begin{bmatrix} -\cos \varphi \operatorname{sen} \theta & -\operatorname{sen} \varphi \cos \theta \\ \cos \varphi \cos \theta & -\operatorname{sen} \varphi \operatorname{sen} \theta \\ 0 & \cos \varphi \end{bmatrix},$$

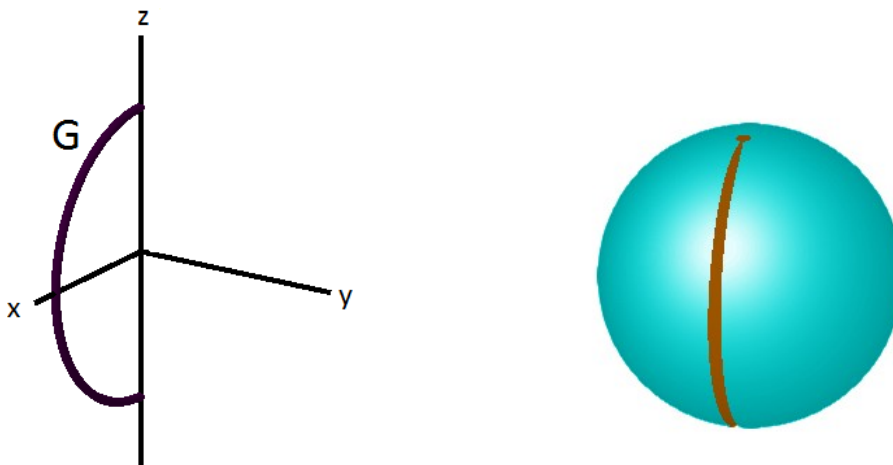
tiene rango = 1 cuando $\cos \varphi = 0$. Esto sugiere tomar la restricción $\Phi|_{W_1}$, siendo:

$$W_1 = \left\{ (\theta, \varphi) : -\frac{\pi}{2} < \varphi < \frac{\pi}{2} \right\} = \mathbb{R} \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),$$

un abierto de $\mathbb{R}_{\theta\varphi}^2$ en el cual $\cos \varphi$ nunca se anula. La nueva parametrización $\Phi|_{W_1}$ así definida es una inmersión, al precio de ya no ser suprayectiva porque su imagen $\Phi(W_1) = S^2 \setminus \{N, S\}$ es la esfera menos los polos norte $N = (0, 0, 1)$ y sur $S = (0, 0, -1)$. Además tampoco es inyectiva, pues cada punto $p \in S^2 \setminus \{N, S\}$ tiene infinitas preimágenes en W_1 . Esto nos sugiere tomar la restricción más pequeña $\Phi|_{W_2}$, donde W_2 es el siguiente abierto de $\mathbb{R}_{\theta\varphi}^2$:

$$W_2 = \left\{ (\theta, \varphi) : 0 < \theta < 2\pi, -\frac{\pi}{2} < \varphi < \frac{\pi}{2} \right\} = (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Ahora $\Phi|_{W_2}$ es una parametrización de máxima calidad: regular e inyectiva. La parte de la esfera parametrizada por $\Phi|_{W_2}$ es el “cascabel” $\Phi(W_2) = S^2 \setminus G$ que resulta de quitarle a la esfera el “meridiano de Greenwich” $G = \{\Phi(0, \varphi) : -\pi/2 \leq \varphi \leq \pi/2\}$.



Al ser imagen continua del intervalo $[-\pi/2, \pi/2]$, el conjunto G es compacto y es cerrado relativo en cualquier conjunto que lo contenga, luego $S^2 \setminus G$ es un abierto relativo de la esfera. Como Φ es biyectiva de W_2 a $S^2 \setminus G$, para cada punto $p \in S^2 \setminus G$ tenemos bien definido el valor paramétrico:

$$(\Phi|_{W_2})^{-1}(p) = (\theta(p), \varphi(p)).$$

Esto define un par de funciones escalares $\theta, \varphi : S^2 \setminus G \rightarrow \mathbb{R}$ que son **funciones coordenadas curvilíneas** en el abierto relativo $S^2 \setminus G$, en el sentido de que cada punto $p \in S^2 \setminus G$ está determinado por los números $\theta(p), \varphi(p)$ y además la función Φ que reconstruye el punto a partir de esos números es \mathcal{C}^∞ y regular.

Se dice que un punto $p \in X$ admite el valor paramétrico (u_1, \dots, u_k) como **coordenadas** (en una parametrización Φ) si es $p = \Phi(u_1, \dots, u_k)$. De esta manera las parametrizaciones proporcionan una **representación numérica de puntos** de la variedad: fijada Φ , especificamos un punto de X cada vez que damos un valor numérico $(u_1, \dots, u_k) \in W$.

Atención. No olvides que no quedan representados los puntos de X que no estén en $\Phi(W)$. Si hay tales puntos, hacen necesarias varias parametrizaciones para cubrir toda X .

Las parametrizaciones también proporcionan **representaciones numéricas de funciones**.

Primer caso: función que sale de la variedad. Dadas una función $f : X \rightarrow \mathbb{R}^m$ y una parametrización $\Phi : W \rightarrow X$, la compuesta

$$g : W \rightarrow \mathbb{R}^m \quad , \quad g(u_1, \dots, u_k) = f \circ \Phi(u_1, \dots, u_k) ,$$

representa *una parte* de la función f : es una representación numérica de $f|_{\Phi(W)}$. Conociendo la “fórmula en coordenadas” $g(u_1, \dots, u_k)$ conocemos el valor $f(p)$ para todo $p \in \Phi(W)$, pero ¡ajo! lo desconocemos para los $p \in X$ que no estén en $\Phi(W)$, si los hay.

Segundo caso: función que llega a la variedad. En este caso tenemos un abierto $U \subseteq \mathbb{R}^m$ y una función $f : U \rightarrow \mathbb{R}^n$ tal que $f(U) \subseteq X$, lo que nos permite considerar a f , de manera alternativa, como una aplicación de U a X .

Supongamos que, de hecho, es $f(U) \subseteq \Phi(W)$ y que Φ es inyectiva. Por razones puramente conjuntistas, existe una única función $g : U \rightarrow W$ tal que $f = \Phi \circ g$. Con más detalle:

$$g(x) \equiv (u_1(x), \dots, u_k(x)) \quad \text{tal que} \quad f(x) \equiv \Phi(u_1(x), \dots, u_k(x)) .$$

Dado cualquier valor $x = (x_1, \dots, x_m) \in U$, lo que hace la “fórmula” $g(x)$ es darnos las coordenadas $(u_1(x), \dots, u_k(x))$ del punto $f(x)$ a partir de los números x_1, \dots, x_m . No es difícil demostrar que si además Φ es regular y \mathcal{C}^s y si f es también \mathcal{C}^s , entonces $g \in \mathcal{C}^s$.

A veces son muy útiles las **parametrizaciones grafo**, que describimos a continuación. Dados un abierto $W \subseteq \mathbb{R}^k$ y una función $\varphi : W \rightarrow \mathbb{R}^{n-k}$ de clase \mathcal{C}^s , la parametrización grafo básica es la siguiente:

$$W \rightarrow \mathbb{R}^n \quad , \quad W \ni u \mapsto (u, \varphi(u)) ,$$

pero se obtienen otras, igual de válidas, aplicando una **matriz de permutación**:

$$P = [\mathbf{e}_{i_1} | \dots | \mathbf{e}_{i_k} | \mathbf{e}_{j_1} | \dots | \mathbf{e}_{j_{n-k}}]_{n \times n} ,$$

matriz ortogonal cuyas columnas son una permutación de la base estándar, dando lugar a:

$$\Phi(u) = P(u, \varphi(u)) = u_1 \mathbf{e}_{i_1} + \dots + u_k \mathbf{e}_{i_k} + \varphi_1(u) \mathbf{e}_{j_1} + \dots + \varphi_{n-k}(u) \mathbf{e}_{j_{n-k}} ,$$

que es regular e inyectiva. La imagen $\Phi(W)$ es un grafo en el que las variables $(x_{j_1}, \dots, x_{j_{n-k}})$ son funciones \mathcal{C}^s de las variables $(x_{i_1}, \dots, x_{i_k})$ con dominio W . Por ejemplo:

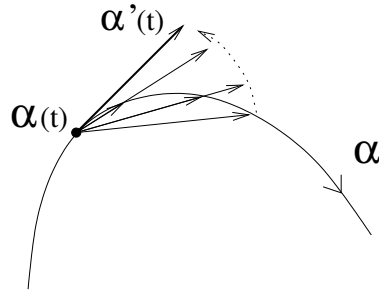
$$\mathbb{R}_{u_1 u_2}^2 \supseteq W \xrightarrow{\Phi} \mathbb{R}^5 \quad , \quad \Phi(u_1, u_2) = (\varphi_1(u_1, u_2), u_1, \varphi_2(u_1, u_2), \varphi_3(u_1, u_2), u_2) .$$

Las parametrizaciones grafo son regulares e inyectivas. Tienen, además, la buena propiedad de que su inversa $\Phi(u) \mapsto u$ viene dada por $\Phi(W) \ni x \mapsto (x_{i_1}, \dots, x_{i_k})$, o sea $\Phi^{-1} = \pi_{i_1 \dots i_k}|_{\Phi(W)}$ es la restricción a $\Phi(W)$ de una **proyección**

$$\pi_{i_1 \dots i_k} : \mathbb{R}^n \rightarrow \mathbb{R}^k \quad , \quad x \mapsto (x_{i_1}, \dots, x_{i_k}) .$$

4.6 Espacios tangentes

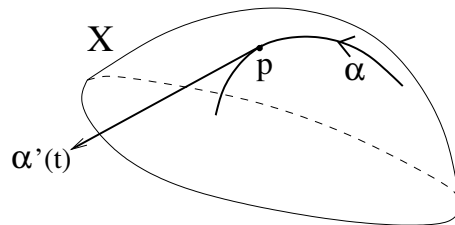
Sean $I \subseteq \mathbb{R}$ un intervalo y $\alpha(t) = (x_1(t), \dots, x_n(t)) : I \rightarrow \mathbb{R}^n$ un camino diferenciable. Sea $t \in I$ un valor paramétrico en el cual $\alpha'(t) \neq \mathbf{0}$. Para algún índice i la derivada $x'_i(t)$ es no nula. Entonces, para h no nulo y pequeño, tenemos $x_i(t+h) \neq x_i(t)$ y los puntos $\alpha(t+h)$, $\alpha(t)$, al ser distintos, determinan una **recta secante** a la curva $\alpha(I)$. Esta recta es paralela al vector $\lambda \cdot (\alpha(t+h) - \alpha(t))$ para cualquier escalar $\lambda \neq 0$. En particular, el vector $\frac{1}{h} \cdot (\alpha(t+h) - \alpha(t))$ da la dirección de la recta secante. Cuando $h \rightarrow 0$ el punto $\alpha(t+h)$ se junta con el punto $\alpha(t)$ y el límite de las rectas secantes es la **recta tangente** a la curva en el punto $\alpha(t)$. Por lo tanto el vector límite $\alpha'(t) = \lim_{h \rightarrow 0} \frac{1}{h} \cdot (\alpha(t+h) - \alpha(t))$ define la dirección de esa recta tangente.



Por esta razón se dice que $\alpha'(t)$ es un vector tangente a la curva $\alpha(I)$ en el punto $\alpha(t)$. Sean ahora una variedad X en \mathbb{R}^n , un punto $p \in X$ y un camino $\alpha : I \rightarrow \mathbb{R}^n$ que cumple las dos condiciones siguientes:

1. Contenido en la variedad: $\alpha(I) \subseteq X$,
2. Pasa por el punto: $\alpha(t) = p$ para algún $t \in I$,

entonces al vector $\alpha'(t)$, tangente a un camino contenido en X , lo consideramos tangente a X ; más aún, lo consideramos **tangente a X en el punto p** .



Definición 109. Dada una variedad X en \mathbb{R}^n y fijado un punto $p \in X$, el **espacio tangente a X en p** es el siguiente conjunto de vectores:

$$T_p X \stackrel{\text{def}}{=} \{ \alpha'(0) : \alpha : I \rightarrow \mathbb{R}^n \text{ camino diferenciable, } 0 \in I, \alpha(I) \subseteq X, \alpha(0) = p \}.$$

Los elementos de este conjunto se llaman **vectores tangentes a X en p** .

Dos comentarios:

(1) Si $\alpha'(0) = \mathbf{0}$ puede que la curva α no tenga una recta tangente bien determinada en $t = 0$ pero, por definición, incluimos el vector nulo en el conjunto $T_p X$. Es decir que aceptamos como elementos de $T_p X$ todas las derivadas $\alpha'(0)$, nulas y no nulas.

(2) Si $\alpha : I \rightarrow \mathbb{R}^n$ es un camino diferenciable contenido en X , entonces para cada $t \in I$ el vector $\alpha'(t)$ es tangente a X en $\alpha(t)$:

$$\alpha(I) \subseteq X \implies \alpha'(t) \in T_{\alpha(t)} X \text{ para todo } t \in I.$$

En efecto, fijado $t \in I$ definimos el nuevo intervalo $J = \{s - t : s \in I\}$ que cumple $0 \in J$, y el nuevo camino $\beta : J \rightarrow \mathbb{R}^n$ dado por $\beta(s) \equiv \alpha(s + t)$. Entonces β es diferenciable y está contenido en X , luego el vector $\beta'(0) = \alpha'(t)$ pertenece a $T_{\beta(0)} X$. Pero $\beta(0) = \alpha(t)$.

Teorema 110. Dada una variedad X en \mathbb{R}^n , cada espacio tangente $T_p X$ es un subespacio vectorial de \mathbb{R}^n y su dimensión como espacio vectorial coincide con la dimensión geométrica de X .

Dados un abierto $U \subseteq \mathbb{R}^m$ y una función diferenciable $f : U \rightarrow \mathbb{R}^n$ con $f(U) \subseteq X$, para cada $a \in U$ el espacio imagen $(df)_a(\mathbb{R}^m)$ está contenido en $T_{f(a)} X$.

Demostración. Sea $k = \dim X$. Fijado un punto $p \in X$, la proposición 105 proporciona un abierto $E \subseteq \mathbb{R}^n$, con $p \in E$, tal que la parte $X \cap E$ es un grafo en el que unas variables $(x_{j_1}, \dots, x_{j_{n-k}})$ son puestas como funciones \mathcal{C}^s de las otras variables $(x_{i_1}, \dots, x_{i_k})$ que a su vez recorren un abierto W de \mathbb{R}^k . Esto equivale a que $X \cap E$ es la imagen $\Phi(W)$ de una parametrización grafo $\Phi(u) = P(u, \varphi(u))$ de las descritas al final del apartado 4.5.

Sean $I \subseteq \mathbb{R}$ un intervalo con $0 \in I$ y $\alpha : I \rightarrow \mathbb{R}^n$ cualquier camino diferenciable contenido en X y tal que $\alpha(0) = p$. Como α es continua y E abierto, existe un intervalo $0 \in J \subseteq I$ tal que $\alpha(J) \subseteq E$. Entonces $\alpha : J \rightarrow X \cap E$ es, en realidad, de la siguiente forma:

$$J \ni t \mapsto \alpha(t) = P(u(t), \varphi(u(t))), \quad (50)$$

donde la función $u(t) : J \rightarrow W \subseteq \mathbb{R}^k$ es simplemente la proyección $\pi_{i_1 \dots i_k} \circ \alpha(t)$ y por lo tanto es diferenciable. Pero la identidad (50) es lo mismo que $\alpha(t) = \Phi \circ u(t)$, de donde:

$$\alpha'(0) = (d\Phi)_{u(0)}(u'(0)).$$

Para todos esos caminos, y las correspondientes funciones $u(t)$, el valor $u^0 = u(0)$ es siempre el mismo: el único $u^0 \in W$ tal que $p = \Phi(u^0)$. Tenemos, pues, un valor paramétrico fijado u^0 tal que todos los caminos α en X con $\alpha(0) = p$ verifican lo siguiente:

$$\alpha'(0) = (d\Phi)_{u^0}(u'(0)),$$

lo que nos lleva a la inclusión $T_p X \subseteq (d\Phi)_{u^0}(\mathbb{R}^k)$.

Si demostramos la inclusión recíproca $T_p X \supseteq (d\Phi)_{u^0}(\mathbb{R}^k)$, quedará probado que $T_p X$ es un espacio vectorial y que su dimensión es igual al rango de $(d\Phi)_{u^0}$, que es k . Pero esto es un caso particular de la segunda afirmación del teorema, que es lo que vamos a probar ahora.

Tenemos, pues, un abierto $U \subseteq \mathbb{R}^m$, una función diferenciable $f : U \rightarrow \mathbb{R}^n$ tal que $f(U) \subseteq X$ y un punto $a \in U$ tal que $f(a) = p$. Queremos probar que $T_p X \supseteq (df)_a(\mathbb{R}^m)$.

Dado cualquier vector $v \in \mathbb{R}^m$ el camino $\beta(t) = a + tv$ está contenido en U para valores pequeños de t . El camino imagen $\alpha(t) = f(a + tv)$ está contenido en X y pasa por p en $t = 0$, luego

$$T_p X \ni \alpha'(0) = (df)_a(v).$$

Como v era cualquier vector de \mathbb{R}^m , queda visto que $T_p X \supseteq (df)_a(\mathbb{R}^m)$. □

Corolario 111. Si $\Psi : W \rightarrow \mathbb{R}^n$ es cualquier parametrización regular para X , entonces para todo $u^0 \in W$ el conjunto $\{\Psi_{u_1}(u^0), \dots, \Psi_{u_k}(u^0)\}$ es una base de $T_{\Psi(u^0)}X$.

Considerando una parametrización grafo $\Phi(u) = (u, \varphi(u))$ y haciendo $L = (d\Phi)_{u^0}$ se tiene:

$$\Phi_{u_1} = (\mathbf{e}_1, \varphi_{u_1}) = (\mathbf{e}_1, L(\mathbf{e}_1)), \dots, \Phi_{u_k} = (\mathbf{e}_k, \varphi_{u_k}) = (\mathbf{e}_k, L(\mathbf{e}_k)),$$

y se deduce que el espacio tangente a $\Phi(W)$ en el punto $p = \Phi(u^0) = (u^0, \varphi(u^0))$ es el conjunto $\{(v, L(v)) : v \in \mathbb{R}^k\}$, o sea el grafo de L .

4.7 Diferencial intrínseca

La segunda parte del teorema 110 nos dice cómo interactúa el espacio tangente con funciones que llegan a la variedad $U \rightarrow X$. Ahora explicaremos cómo interactúa con funciones que salen de la variedad $X \rightarrow \mathbb{R}^m$.

Definición 112. Sean X una variedad en \mathbb{R}^n y $p \in X$. Una función $f : X \rightarrow \mathbb{R}^m$ es **diferenciable en p (como función en X)** si existen funciones lineales $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tales que:

$$\text{la función } X \rightarrow \mathbb{R}^m \text{ dada por } q \mapsto f(q) - f(p) - L(q - p) \text{ es un } o(\|q - p\|). \quad (51)$$

Como todas las normas en \mathbb{R}^n son equivalentes, para todas ellas es la misma la clase de funciones que son un $o(\|q - p\|)$.

Veamos, con un ejemplo, que la función lineal L no es única. Sean $X = \{y = x^2\} \subset \mathbb{R}^2$ y $f : X \rightarrow \mathbb{R}$ dada por $f(x, y) = x e^y$. Nos fijamos en el punto $p = (0, 0) \in X$ y para cada constante c consideramos la función lineal $L_c(x, y) = x + cy$. Describiendo el punto general de X como (x, x^2) , calculamos:

$$\begin{aligned} f(x, x^2) - f(0, 0) - L_c((x, x^2) - (0, 0)) &= f(x, x^2) - 0 - L_c(x, x^2) = \\ &= x e^{x^2} - (x + c x^2) = x(e^{x^2} - 1) - c x^2 = x O(x^2) + O(x^2) = O(|x|^3) + O(x^2) = \\ &= O(x^2) = o(|x|) = o(\|(x, x^2)\|) = o(\|(x, x^2) - (0, 0)\|). \end{aligned}$$

Vemos que, según la definición 112, la función f es diferenciable en p como función $X \rightarrow \mathbb{R}$, valiendo para ello cualquiera de las funciones lineales L_c . Se comprueba fácilmente que no vale ninguna otra: si el coeficiente de x en L no es 1 entonces $f(x, x^2) - L(x, x^2) \neq o(\|(x, x^2)\|)$. La función lineal L está, pues, *parcialmente determinada*: uno de sus coeficientes está fijo y el otro es libre. Otra manera de decirlo es que el valor $L(v)$ es único para los vectores particulares $v = (a, 0)$; pero éstos son, precisamente, los vectores tangentes a X en el punto p .

Teorema-definición 113. Si f es diferenciable en un punto $p \in X$ como función $X \rightarrow \mathbb{R}^m$, entonces todas las funciones lineales L que cumplen (51) tienen la misma restricción $L|_{T_p X}$ al espacio tangente.

Llamamos **diferencial intrínseca de f en p** a la función lineal $(df)_p : T_p X \rightarrow \mathbb{R}^m$ definida por una cualquiera de esas restricciones.

Demostración. Elegimos una tal L y un vector cualquiera $v \in T_p X$. Por definición del espacio tangente, existe un camino diferenciable $\alpha : I \rightarrow \mathbb{R}^n$ con $\alpha(I) \subseteq X$, $\alpha(0) = p$ y $\alpha'(0) = v$. Aplicamos (51) y obtenemos:

$$\begin{aligned} f(\alpha(t)) - f(\alpha(0)) &= \\ &= f(\alpha(t)) - f(p) = L(\alpha(t) - p) + o(\|\alpha(t) - p\|) = \\ &= L(\alpha(t) - \alpha(0)) + o(\|\alpha(t) - \alpha(0)\|), \end{aligned}$$

pero también tenemos $\alpha(t) - \alpha(0) = t \alpha'(0) + o(|t|) = t v + o(|t|) = O(|t|)$, luego:

$$\begin{aligned} f(\alpha(t)) - f(\alpha(0)) &= L(t v + o(|t|)) + o(O(|t|)) = \\ &= t L(v) + L(o(|t|)) + o(|t|) = t L(v) + o(|t|), \end{aligned}$$

de donde:

$$\lim_{t \rightarrow 0} \frac{f(\alpha(t)) - f(\alpha(0))}{t} = \lim_{t \rightarrow 0} \frac{t L(v) + o(|t|)}{t} = L(v),$$

y llegamos a la igualdad $\left. \frac{d}{dt} \right|_{t=0} f(\alpha(t)) = L(v)$, en cuyo miembro de la izquierda no aparece L . Esto demuestra que, cuando $v \in T_p X$, el valor $L(v)$ sólo depende de los datos X, p, f, v y no de la L que hayamos elegido cumpliendo (51). \square

Esta demostración nos da una fórmula para la diferencial intrínseca:

$$(df)_p(\alpha'(0)) = \left. \frac{d}{dt} \right|_{t=0} f(\alpha(t)), \quad (52)$$

donde el camino α tiene que estar contenido en X y pasar por p cuando $t = 0$.

Formalmente, la fórmula (52) es igual que la regla de la cadena para caminos del apartado 2.3.

El espacio tangente en un punto es el dominio natural de la diferencial intrínseca de una función que sale de una variedad. Dicha diferencial es un objeto que satisface una regla de la cadena para caminos en la variedad.

Ahora nos planteamos cómo calcular la diferencial intrínseca. Vamos a tratar dos casos.

Primer caso: tenemos parámetros explícitos. Más en concreto, hay un abierto $W \subseteq \mathbb{R}^k$, una parametrización $\Phi : W \rightarrow X$ y un valor $a \in W$ tal que $\Phi(a) = p$ y $(d\Phi)_a$ tiene rango $k = \dim X$. El corolario 111 nos dice que $\{\Phi_{u_1}(a), \dots, \Phi_{u_k}(a)\}$ es una base de $T_p X$ y la igualdad:

$$(df)_p(c_1 \Phi_{u_1}(a) + \dots + c_k \Phi_{u_k}(a)) = c_1 (df)_p(\Phi_{u_1}(a)) + \dots + c_k (df)_p(\Phi_{u_k}(a)),$$

nos dice que basta calcular los valores $(df)_p(\Phi_{u_1}(a)), \dots, (df)_p(\Phi_{u_k}(a))$ para conocer $(df)_p(v)$ para el vector general $v \in T_p X$. Dado $1 \leq i \leq k$, el camino $\beta(t) = a + t e_i$ está en W para t pequeño y $\alpha(t) = \Phi \circ \beta(t)$ es un camino en X con $\alpha(0) = p$ y $\alpha'(0) = \Phi_{u_i}(a)$. Aplicando (52):

$$(df)_p(\Phi_{u_i}) = (df)_p(\alpha'(0)) = \left. \frac{d}{dt} \right|_{t=0} f \circ \alpha(t) = \left. \frac{d}{dt} \right|_{t=0} (f \circ \Phi) \circ \beta(t) = (f \circ \Phi)_{u_i}(a).$$

La diferencial intrínseca es la única función lineal $(df)_p : T_p X \rightarrow \mathbb{R}^m$ con el siguiente efecto sobre la base del corolario 111: $\Phi_{u_1}(a) \mapsto (f \circ \Phi)_{u_1}(a), \dots, \Phi_{u_k}(a) \mapsto (f \circ \Phi)_{u_k}(a)$.

Segundo caso: f es restricción a X de una función de n variables. Más en concreto, tenemos un abierto $U \subseteq \mathbb{R}^n$ que contiene a X , una función diferenciable $F : U \rightarrow \mathbb{R}^m$ y definimos f como la restricción $f = F|_X$. Dados $p \in X$ y $v \in T_p X$, sabemos que existen caminos α , contenidos en X , con $\alpha(0) = p$ y $\alpha'(0) = v$. Con uno de estos caminos calculamos:

$$(df)_p(v) = (df)_p(\alpha'(0)) = \left. \frac{d}{dt} \right|_{t=0} f \circ \alpha(t) = \left. \frac{d}{dt} \right|_{t=0} F \circ \alpha(t) = (dF)_p(v),$$

y resulta que $(df)_p$ es, simplemente, la restricción a $T_p X$ de la diferencial de F .

$$(df)_p = (dF)_p|_{T_p X}.$$

Esto es conveniente para ejemplos como el siguiente:

$$X = \{(x, y) \in \mathbb{R}^2 : x + e^x + 2y + \cos y = 2\},$$

que es un grafo $y = y(x)$ y también un grafo $x = x(y)$, pero ninguna de esas dos funciones es elemental y no tenemos una parametrización explícita para X . El punto $p = (0, 0)$ está en X y, derivando implícitamente: $1 + e^x + (2 - \sin y(x)) y'(x) = 0$, sacamos $y'(0) = -1$. La recta tangente $T_p X$ está generada por el vector $(1, -1)$. Para cualquier función diferenciable $F(x, y)$, definida en un abierto del plano que contenga a X , y la correspondiente función $f = F|_X$, tenemos entonces: $(df)_p((a, -a)) = (dF)_p((a, -a)) = a F_x(0, 0) - a F_y(0, 0) = DF_{(0,0)} \begin{bmatrix} a \\ -a \end{bmatrix}$.

4.8 Espacio normal

Definición 114. Sea X una variedad de dimensión k en \mathbb{R}^n y sea $p \in X$. El **espacio normal a X en p** es el complemento ortogonal de $T_p X$, es decir

$$\{w \in \mathbb{R}^n : w \cdot v = 0 \text{ para todo } v \in T_p X\}.$$

Un vector de \mathbb{R}^n es **normal a X en p** si pertenece al espacio normal. La **codimensión de X** es la dimensión $n - k$ de sus espacios normales.

Las hipersuperficies, definidas en el apartado 4.2, son las variedades de codimensión 1 en \mathbb{R}^n . El teorema 104 del apartado 4.3 dice que si una variedad está definida por un sistema de m ecuaciones escalares $f(x) = b$, y los gradientes de las ecuaciones son linealmente independientes en cada punto de la variedad, entonces la codimensión de la variedad es igual al número de ecuaciones del sistema.

Podemos entender la dimensión geométrica como una “medida del grosor de X ” y la codimensión como una “medida de su delgadez”: cada vez que se añade una ecuación escalar (con gradiente independiente de los gradientes de las ecuaciones previas) al sistema $f(x) = b$, el conjunto de soluciones se hace más delgado.

Lema 115. Sean X una variedad en \mathbb{R}^n y F una función diferenciable definida en un abierto de \mathbb{R}^n que contiene a X . Si F es constante en X entonces para cada punto $p \in X$ el gradiente ∇F_p es normal a X en p .

Demostración. Sea $v \in T_p X$. Existe un camino α contenido en X , con $\alpha(0) = p$ y $\alpha'(0) = v$. Calculamos:

$$v \cdot \nabla_p F = (DF_p)v = \left. \frac{d}{dt} \right|_{t=0} F \circ \alpha(t) = \left. \frac{d \text{cte}}{dt} \right|_0 = 0.$$

□

Proposición 116. Sean $U_0 \subseteq \mathbb{R}^n$ un abierto y $f = (f_1, \dots, f_m) : U_0 \rightarrow \mathbb{R}^m$ una función de clase C^s que es regular en todos los puntos del conjunto $X = \{x \in U_0 : f(x) = b\}$, que por lo tanto es una variedad en \mathbb{R}^n . Para cada punto $p \in X$, los gradientes $\nabla f_1(p), \dots, \nabla f_m(p)$ forman una base del espacio normal a X en p .

Demostración. Sea $p \in X$. Cada función f_i es constante en X , luego cada gradiente $\nabla f_i(p)$ es normal a X en p . La hipótesis de regularidad dice que $\nabla f_1(p), \dots, \nabla f_m(p)$ son linealmente independientes, luego son base de un subespacio \mathbb{V} de dimensión m del espacio normal a X en p . Pero $m = n - \dim X = n - \dim T_p X$ es la dimensión del espacio normal, luego \mathbb{V} es todo el espacio normal. □

Corolario 117. En las condiciones de la proposición 116, para cada punto $p \in X$ se tiene:

$$T_p X = \{v \in \mathbb{R}^n : v \cdot \nabla f_1(p) = \dots = v \cdot \nabla f_m(p) = 0\}.$$

Veamos un ejemplo. Sea $X \subset \mathbb{R}^4$ el conjunto dado por $e^{x_1}(x_2 + x_3 + x_4) + x_2^3 x_3 e^{x_4} = e^2$. El gradiente de la ecuación:

$$(e^{x_1}(x_2 + x_3 + x_4), e^{x_1} + 3x_2^2 x_3 e^{x_4}, e^{x_1} + x_2^3 e^{x_4}, e^{x_1} + x_2^3 x_3 e^{x_4}),$$

no se anula en ningún punto de X . En efecto, en un punto $x \in X$, donde se anule la primera componente del gradiente, tiene que cumplirse $x_2^3 x_3 e^{x_4} = e^2$, con lo cual la cuarta componente del gradiente $e^{x_1} + e^2$ es positiva. Luego X es una variedad de dimensión $4 - 1 = 3$ y clase C^∞ en \mathbb{R}^4 . Dado $p = (2, -1, -1, 2) \in X$, nos planteamos hallar una base de $T_p X$.

Primer método: derivación implícita y corolario 111. El gradiente de la ecuación en p es $(0, -2e^2, 0, 2e^2)$ y, por ejemplo, hay un abierto $E \subset \mathbb{R}^4$, con $p \in E$, tal que $X \cap E$ es el grafo $\{x_2 = \varphi(x_1, x_3, x_4)\}$ de una función φ de clase C^∞ definida en un abierto $W \subset \mathbb{R}_{x_1 x_3 x_4}^3$ que contiene el punto $u^0 = (p_1, p_3, p_4) = (2, -1, 2)$. Además tenemos el dato $\varphi(u^0) = p_2 = -1$, que nos va a permitir calcular implícitamente las derivadas de φ en u^0 . En W se cumple la identidad:

$$e^{x_1} (\varphi(x_1, x_3, x_4) + x_3 + x_4) + \varphi(x_1, x_3, x_4)^3 x_3 e^{x_4} \equiv e^2,$$

que al derivarla respecto de x_1, x_3 y x_4 produce la siguiente identidad vectorial:

$$(e^{x_1} + 3\varphi^2 x_3 e^{x_4}) (\varphi_{x_1}, \varphi_{x_3}, \varphi_{x_4}) + ((\varphi + x_3 + x_4) e^{x_1}, e^{x_1} + \varphi^3 e^{x_4}, e^{x_1} + \varphi^3 x_3 e^{x_4}) \equiv (0, 0, 0).$$

El dato $\varphi(u^0) = -1$ nos permite evaluar esta identidad en u^0 , obteniéndose la igualdad vectorial $-2e^2 (\varphi_{x_1}, \varphi_{x_3}, \varphi_{x_4})_{u^0} + (0, 0, 2e^2) = (0, 0, 0)$ de la que se despeja:

$$(\varphi_{x_1}, \varphi_{x_3}, \varphi_{x_4})_{u^0} = (0, 0, 1). \quad (53)$$

La correspondiente parametrización grafo es:

$$\Phi(u_1, u_3, u_4) = (u_1, \varphi(u_1, u_3, u_4), u_3, u_4) \quad \text{con} \quad \Phi(u^0) = p,$$

y, utilizando el corolario 111 y los datos (53), una base de $T_p X$ es:

$$\{ \Phi_{u_1}(u^0), \Phi_{u_3}(u^0), \Phi_{u_4}(u^0) \} = \{ (1, 0, 0, 0), (0, 0, 1, 0), (0, 1, 0, 1) \}.$$

Segundo método: corolario 117. Un vector $v = (v_1, v_2, v_3, v_4)$ es tangente a X en p si y sólo si es ortogonal al gradiente en p de la ecuación: $0 = v \cdot (0, -2e^2, 0, 2e^2) = 2e^2(v_4 - v_3)$, de donde $T_p X = \{v_3 = v_4\}$. Esto vuelve a darnos la base anterior, pero con muchos menos cálculos.

Veamos otro ejemplo, pero ahora sólo usaremos el corolario 117, que hemos visto es más limpio. Sea ahora X el conjunto definido por el sistema (48) del apartado 4.3. En el punto $a' = (0, 0, 0, 1) \in X$ vimos que la jacobiana del sistema tiene las dos filas linealmente independientes, por lo tanto hay un abierto $E \subset \mathbb{R}^4$, con $a' \in E$, tal que $X \cap E$ es una variedad de dimensión $4 - 2 = 2$ y clase C^∞ , en \mathbb{R}^4 . Un vector $v \in \mathbb{R}^4$ es tangente a $X \cap E$ en a' si y sólo si es ortogonal a los gradientes en a' de las dos ecuaciones de (48), es decir ortogonal a las dos filas de la jacobiana en a' :

$$\begin{bmatrix} 0 & 8 & 0 & 0 \\ -1 & 5 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Las soluciones de este sistema vienen dadas por $v_1 = v_2 = 0$, luego $T_{a'}(X \cap E) = \{(0, 0)\} \times \mathbb{R}^2$.

Observación. El segundo ejemplo deja claro que si X viene dada por un sistema de ecuaciones $f(x) = b$, con los gradientes de las ecuaciones ∇f_i linealmente independientes en cada punto de X , entonces para todo punto $p \in X$ el espacio tangente $T_p X$ puede calcularse como un núcleo: $T_p X = \ker(df)_p$

4.9 Máximos y mínimos condicionados

Supongamos que queremos encontrar el punto más cercano al origen en la elipse

$$X = \{(x, y) \in \mathbb{R}^2 : x^2 - xy + y^2 = 1\}.$$

Esto supone hallar el mínimo en X de la función $F(x, y) = \sqrt{x^2 + y^2}$. Es habitual llamar a ese valor el **mínimo de F sujeto a la condición $x^2 - xy + y^2 = 1$** , y por lo tanto es un ejemplo de **mínimo condicionado**. También podemos buscar el punto más lejano del origen en dicha elipse, es decir el **máximo de F sujeto a la condición $x^2 - xy + y^2 = 1$** , que es un ejemplo de **máximo condicionado**; dicho punto será también el más lejano del origen en la región R limitada por la elipse, es decir $R = \{(x, y) : x^2 - xy + y^2 \leq 1\}$.

Es de señalar que F tiene un mínimo en todo su dominio: $F(0, 0) = 0$, pero que en X va a tener un mínimo diferente, estrictamente positivo. Por otra parte F no tiene máximo en todo su dominio, pero en X y en R sí va a alcanzar un máximo finito.

En general, un problema de mínimo condicionado pide hallar el mínimo de una función F en un conjunto X definido por unas cuantas *condiciones*, cada una de las cuales será una ecuación o una desigualdad. Análogamente un máximo condicionado. El máximo y el mínimo condicionados, así como su existencia, dependerán de F y de las condiciones impuestas.

Una situación frecuente es la siguiente. Tenemos un abierto $U \subseteq \mathbb{R}^n$, una función escalar diferenciable $F : U \rightarrow \mathbb{R}$, una variedad X contenida en U , buscamos el máximo (o el mínimo, o ambos) de F en X y de hecho nos preguntamos por el punto o puntos de X donde se alcanza ese máximo (o el mínimo).

Proposición 118. Para que $p \in X$ sea un punto de máximo local o de mínimo local de la función $f = F|_X$ es necesario (no suficiente) que p sea un **punto crítico** de la función $f : X \rightarrow \mathbb{R}$, es decir que se anule la diferencial intrínseca: $(df)_p = 0$.

Demostración. Vamos a demostrar que si p no es crítico para f entonces no puede ser ni máximo local ni mínimo local para f .

Suponemos, pues $(df)_p \neq 0$. Habrá algún vector $v \in T_p X$ con $(df)_p(v) \neq 0$ y tomando un camino α , contenido en X , con $\alpha(0) = p$ y $v = \alpha'(0)$, tendremos:

$$c \stackrel{\text{def}}{=} \left. \frac{d}{dt} \right|_{t=0} f(\alpha(t)) = (df)_p(\alpha'(0)) \neq 0.$$

Si $c > 0$ entonces será $f(\alpha(t)) > f(p)$ para t positivo y pequeño y será $f(\alpha(t)) < f(p)$ para t negativo y pequeño. Si $c < 0$ se invierten esas desigualdades. En ambos casos habrá puntos $q = \alpha(t)$, situados en X y arbitrariamente cerca de p , algunos con $f(q) > f(p)$ y otros con $f(q) < f(p)$, con lo cual p no será ni máximo local ni mínimo local para f . \square

Por otra parte, en el apartado 4.7 hemos visto que la diferencial intrínseca de $f = F|_X$ es la restricción de $(dF)_p$ al espacio tangente $T_p X$. Se deduce que las condiciones siguientes son equivalentes:

- p es un punto crítico de f .
- Para todo $v \in T_p X$ se tiene $0 = (df)_p(v) = (dF)_p(v) = \nabla F(p) \cdot v$.
- El vector $\nabla F(p)$ es normal a X en p .

Esto, junto con la proposición 116, nos lleva al siguiente resultado.

Teorema-definición 119. Sean un abierto $U \subseteq \mathbb{R}^n$, una función diferenciable $F : U \rightarrow \mathbb{R}$, un conjunto X definido por un sistema de ecuaciones

$$X = \{x \in U : f(x) = (f_1, \dots, f_m)(x) = b\},$$

cuyos gradientes son linealmente independientes en todo punto de X , con lo cual X es una variedad en \mathbb{R}^n . Entonces un punto $p \in X$ es un punto crítico de $f = F|_X$ si y sólo si existen escalares $\lambda_1, \dots, \lambda_m$ tales que:

$$\nabla F(p) = \lambda_1 \nabla f_1(p) + \dots + \lambda_m \nabla f_m(p). \quad (54)$$

Además esos escalares son únicos y se los llama **multiplicadores de Lagrange**.

Combinando la proposición 118 y el teorema 119, llegamos a lo siguiente:

Para que un punto p realice el máximo o el mínimo de F en X es necesario (no suficiente) que cumpla las siguientes ecuaciones:

$$f(x) = p \quad , \quad \nabla F(p) = \lambda_1 \nabla f_1(p) + \dots + \lambda_m \nabla f_m(p) ,$$

para ciertos multiplicadores de Lagrange $\lambda_1, \dots, \lambda_m$.

Vamos a aplicar esto a la búsqueda del punto de la elipse $X = \{x^2 + xy + y^2 = 1\}$ más cercano al origen y el más alejado del origen. esos puntos existen porque la distancia al origen $\sqrt{x^2 + y^2}$ es continua y la elipse es compacta. Minimizar (o maximizar) la distancia equivale a hacerlo con su cuadrado $F^2 = x^2 + y^2$, que es una función más sencilla y además suave en todo el plano. Ahora sabemos que esos puntos de máximo y mínimo satisfacen las igualdades:

$$x^2 - xy + y^2 = 1 \quad , \quad \nabla(x^2 + y^2) = \lambda \nabla(x^2 - xy + y^2) ,$$

donde el escalar λ depende del punto. Esto nos da un sistema:

$$\left. \begin{aligned} x^2 - xy + y^2 &= 1 \\ 2x &= 2\lambda x - \lambda y \\ 2y &= -\lambda x + 2\lambda y \end{aligned} \right\}$$

Enseguida se ve que no hay soluciones con $x = 0$, pues la primera ecuación daría $y = \pm 1$ y entonces la segunda y la tercera se convierten en $(0, y) = \lambda(-y, 2y)$, que es incompatible. Del mismo modo se ve que no hay soluciones con $y = 0$. Esto permite dividir la segunda ecuación por x y la tercera por y , lo cual da: $2 = \lambda \left(2 - \frac{y}{x}\right) = \lambda \left(2 - \frac{x}{y}\right)$, de donde $\lambda \neq 0$ y $\frac{x}{y} = \frac{y}{x}$, luego $y = \pm x$.

En la primera ecuación, el caso $y = x$ da los puntos $p^\pm = \pm(1, 1)$ y el caso $y = -x$ da los puntos $q^\pm = \pm(1, -1)/\sqrt{3}$. En cada uno de esos puntos hallamos el multiplicador de Lagrange y la distancia al origen:

$$\text{en los puntos } p^\pm \text{ es } \lambda = 2 \text{ y } F(p^\pm) = \sqrt{2},$$

$$\text{en los puntos } q^\pm \text{ es } \lambda = 2/3 \text{ y } F(q^\pm) = \sqrt{2/3},$$

Concluimos que los puntos de la elipse más cercanos al origen son los q^\pm , con distancia $\sqrt{2/3}$, y los más lejanos del origen son los p^\pm con distancia $\sqrt{2}$.

Ahora vamos a hallar un máximo y un mínimo en una región definida por una *desigualdad*. Buscamos el máximo y el mínimo de la función $F(x, y) = (x - 1)^2 + 4y^2$, así como los puntos donde se alcanzan, en la región $R = \{x^2 + y^2 \leq 3\}$. Esos puntos existen porque R es compacta. Separamos la región R en su interior, que es la bola abierta $U = \{x^2 + y^2 < 3\}$, y su frontera que es la circunferencia $X = \{x^2 + y^2 = 3\}$.

Para resolver el problema buscamos puntos de máximo y mínimo en U (si es que los hay), puntos de máximo y mínimo en X (que seguro que los hay, porque X es compacta) y comparamos valores de f en los puntos resultantes.

Un máximo o un mínimo de F en el interior, si existe, debe ser crítico para F , es decir que en él debe anularse el gradiente $\nabla F = (2x - 2, 8y)$. El único punto cumpliendo eso es el $p_0 = (1, 0)$, que guardamos para luego.

Un punto de máximo o mínimo de $F|_X$ tiene que satisfacer:

$$x^2 + y^2 = 3 \quad , \quad (2x - 2, 8y) = \lambda(2x, 2y) .$$

Puesto como sistema de ecuaciones escalares:

$$\left. \begin{aligned} x^2 + y^2 &= 3 \\ x - 1 &= \lambda x \\ 4y &= \lambda y \end{aligned} \right\}$$

Las soluciones con $y = 0$ son los puntos $(\pm\sqrt{3}, 0)$. Las soluciones con $y \neq 0$ tienen forzosamente $\lambda = 4$ y $x = -1/3$, luego son los puntos $(-1/3, \pm\sqrt{26}/3)$. Evaluamos:

$$F(p_0) = 0,$$

$$\text{en } p_1 = (\sqrt{3}, 0) \text{ es } \lambda = 1 - (1/\sqrt{3}) \text{ y } F(p_1) = (\sqrt{3} - 1)^2 \approx 0'53,$$

$$\text{en } p_2 = (-\sqrt{3}, 0) \text{ es } \lambda = 1 + (1/\sqrt{3}) \text{ y } F(p_2) = (\sqrt{3} + 1)^2 \approx 7'46,$$

$$\text{en } p^\pm = (-1/3, \pm\sqrt{26}/3) \text{ es } \lambda = 4 \text{ y } F(p^\pm) = 14/3 \approx 4'66.$$

Vemos que 0 es el mínimo de F en toda la región y se alcanza solamente en el punto $p_0 = (1, 0)$, que está en el interior. El máximo de F en la región es $(\sqrt{3} + 1)^2$ y sólo se alcanza en el punto $p_2 = (-\sqrt{3}, 0)$, que está en la frontera. El mínimo de F en la frontera es $(\sqrt{3} - 1)^2$, mayor que el mínimo en el interior, y sólo se alcanza en el punto $p_1 = (\sqrt{3}, 0)$. Los puntos p^\pm son puntos críticos de $F|_X$, pero no son máximos ni mínimos de $F|_X$.

4.10 Hessiana intrínseca

En este apartado las funciones y las variedades se suponen al menos de clase \mathcal{C}^2 .

Sean un abierto $U \subseteq \mathbb{R}^n$ y una función $f : U \rightarrow \mathbb{R}$. La forma Hessiana de f en un punto $p \in U$, tal como la hemos definido en el apartado 2.8, se obtiene al derivar f dos veces a lo largo de pequeños segmentos de recta afín:

$$\text{Hess}(f)_p(v) = \left. \frac{d^2}{dt^2} \right|_{t=0} f(p + tv).$$

No podemos hacer exactamente lo mismo con una función definida en una variedad X en \mathbb{R}^n , porque puede no haber, o haber muy pocos, segmentos de recta contenidos en X .

Por ejemplo, se comprueba que la esfera $\{x^2 + y^2 + z^2 = 1\}$ corta a cada recta afín de \mathbb{R}^3 a lo más en dos puntos. Como todo segmento de recta tiene (aunque sea muy corto) infinitos puntos, la esfera no contiene ningún segmento de recta.

En su lugar, vamos a derivar dos veces a lo largo de caminos contenidos en X . La función vendrá dada como una restricción $f = F|_X$ donde F está definida en un abierto $U \subseteq \mathbb{R}^n$ que contiene a X . Para cualquier camino $\alpha(t) = (x_1(t), \dots, x_n(t))$ contenido en X , y por lo tanto en U , un cálculo fácil nos da lo siguiente:

$$\frac{d^2}{dt^2} f(\alpha(t)) = \frac{d^2}{dt^2} F(\alpha(t)) = \sum_{1 \leq i, j \leq n} x'_i(t) x'_j(t) F_{x_i x_j}(\alpha(t)) + \sum_{i=1}^n x''_i(t) F_{x_i}(\alpha(t)),$$

es decir:

$$(F \circ \alpha)''(t) = \text{Hess}(F)_{\alpha(t)}(\alpha'(t)) + \alpha''(t) \cdot \nabla F(\alpha(t)). \quad (55)$$

El término $\alpha''(t) \cdot \nabla F(\alpha(t))$ no apareció en el apartado 2.8 porque los caminos de la forma $\alpha(t) = p + tv$ tienen $\alpha''(t) \equiv \mathbf{0}$. Pero un camino contenido en la variedad X estará obligado, casi siempre, a tener α'' no nula.

Dado un punto $p \in X$, lo que vamos a hacer ahora es ajustar la función que extiende f al abierto U , es decir cambiar F por otra función $G : U \rightarrow \mathbb{R}$ que también cumpla $f = G|_X$ y además tenga $\nabla G(p) = \mathbf{0}$. Como $(df)_p = (dG)_p|_{T_p X}$, para que una tal G pueda existir es necesario que sea $(df)_p = 0$, es decir que p tiene que ser un punto crítico de $f : X \rightarrow \mathbb{R}$.

La Hessiana intrínseca de una función definida en una variedad sólo se define en los puntos críticos de dicha función.

Es el precio que pagamos por no tener segmentos de recta en X .

Suponemos X definida por un sistema de m ecuaciones escalares:

$$X = \{x \in U : h_1(x) = b_1, \dots, h_m(x) = b_m\},$$

con $h = (h_1, \dots, h_m) : U \rightarrow \mathbb{R}^m$ al menos de clase \mathcal{C}^2 y los gradientes $\nabla h_1, \dots, \nabla h_m$ linealmente independientes en cada punto de X . Entonces sabemos que p es un punto crítico de $f = F|_X$ si y sólo si $\nabla F(p) = \lambda_1 \nabla h_1(p) + \dots + \lambda_m \nabla h_m(p)$, para unos multiplicadores de Lagrange $\lambda_1, \dots, \lambda_m$.

Partiendo de un punto $p \in X$ crítico para $f = F|_X$, calculamos los multiplicadores de Lagrange en este punto y con ellos formamos la función:

$$G_0 = F - \lambda_1 (h_1 - b_1) - \dots - \lambda_m (h_m - b_m).$$

Como cada una de las funciones $f_j - b_j$ es idénticamente nula en X , tenemos $G_0|_X = F|_X = f$. Además

$$\nabla G_0(p) = \nabla F(p) - \lambda_1 \nabla h_1(p) - \dots - \lambda_m \nabla h_m(p) = \mathbf{0}.$$

Teorema-definición 120. Dada $f : X \rightarrow \mathbb{R}$ y el punto $p \in X$ crítico para f , para cualquier camino $\alpha(t) : I \rightarrow \mathbb{R}^n$, contenido en X y con $\alpha(0) = p$, el número $(f \circ \alpha)''(0)$ depende solamente del vector $v = \alpha'(0) \in T_p X$ definiendo así una función escalar $Q(v) = (f \circ \alpha)''(0)$. Esta función $Q : T_p X \rightarrow \mathbb{R}$ resulta ser una forma cuadrática en el espacio $T_p X$ y la llamamos **forma Hessiana intrínseca de f en el punto crítico p** .

Demostración.

Hacemos uso de la función G_0 que acabamos de describir, que cumple $f = G_0|_X$ y $\nabla G_0(p) = \mathbf{0}$. Aplicamos a G_0 la fórmula (55) y obtenemos:

$$\begin{aligned} (f \circ \alpha)''(0) &= (G_0 \circ \alpha)''(0) = \text{Hess}(G_0)_{\alpha(0)}(\alpha'(0)) + \alpha''(0) \cdot \nabla G_0(\alpha(0)) = \\ &= \text{Hess}(G_0)_p(v) + \alpha''(0) \cdot \nabla G_0(p) = \text{Hess}(G_0)_p(v) + \alpha''(0) \cdot \mathbf{0}, \end{aligned}$$

y el resultado no depende de $\alpha''(0)$ porque este vector está multiplicado por $\mathbf{0}$. Llegamos a:

$$(\alpha \subset X \text{ y } \alpha(0) = p \text{ y } \alpha'(0) = v) \implies (f \circ \alpha)''(0) = \overset{\circ}{\text{Hess}}(G_0)_p(v),$$

y, efectivamente, fijados f y el punto crítico p el resultado sólo depende de $v \in T_p X$ y es una forma cuadrática actuando sobre v . \square

Tenemos, además, una fórmula para la Hessiana intrínseca: $Q = \text{Hess}(G_0)_p|_{T_p X}$, pero vamos a hacerle una simplificación. Definimos la función:

$$G = F - \lambda_1 h_1 - \dots - \lambda_m h_m,$$

que es tal que $G_0 = G + \lambda_1 b_1 + \dots + \lambda_m b_m = G + \text{cte}$. Entonces:

$$\text{Hess}(G_0)_p = \text{Hess}(G)_p = \text{Hess}(F)_p - \lambda_1 \text{Hess}(h_1)_p - \dots - \lambda_m \text{Hess}(h_m)_p.$$

Finalmente, pues, la fórmula para la Hessiana intrínseca de f en el punto crítico p es:

$$Q = \left(\text{Hess}(F)_p - \lambda_1 \text{Hess}(h_1)_p - \dots - \lambda_m \text{Hess}(h_m)_p \right) \Big|_{T_p X}$$

y es una forma cuadrática $Q : T_p X \rightarrow \mathbb{R}$ tal que $(f \circ \alpha)''(0) = Q(\alpha'(0))$ para todo camino α contenido en X y tal que $\alpha(0) = p$.

Lema 121. Si la Hessiana intrínseca toma un valor negativo entonces el punto crítico p no es un mínimo local de f . Si la Hessiana intrínseca toma un valor positivo entonces p no es un máximo local de f .

Corolario 122. Para que el punto crítico p sea mínimo local de f es necesario (no suficiente) que la hessiana intrínseca sea semidefinida positiva. Para que p sea máximo local de f es necesario (no suficiente) que la Hessiana intrínseca sea semidefinida negativa.

Si la Hessiana intrínseca es indefinida (degenerada o no) entonces p no es ni máximo local ni mínimo local de f .

Se demuestran igual que el lema 72 y el corolario 73 del apartado 2.9.

Teorema 123. Para que el punto crítico p sea un mínimo local estricto de f es suficiente (no necesario) que la Hessiana intrínseca sea definida positiva. Para que p sea un máximo local estricto de f es suficiente (no necesario) que la Hessiana intrínseca sea definida negativa.

Demostración. Tomamos un abierto E de \mathbb{R}^n contenido en el dominio U de F , con $p \in E$ y tal que la parte $X \cap E$ sea un grafo donde $n-k$ variables sean puestas como funciones \mathcal{C}^s de las otras k variables, que a su vez recorren un abierto $W \subseteq \mathbb{R}^k$. La correspondiente parametrización grafo $\Phi : W \rightarrow \mathbb{R}^n$ es biyectiva de W a $X \cap E$. La función $f \circ \Phi : W \rightarrow \mathbb{R}$ es de clase \mathcal{C}^s porque se la puede escribir como $F \circ \Phi$, compuesta de dos funciones \mathcal{C}^s con el abierto E como dominio intermedio.

Dado el punto $a \in W$ tal que $p = \Phi(a)$, la diferencial $(d\Phi)_a$ es un isomorfismo de \mathbb{R}^k a T_pX . El punto a es crítico para $f \circ \Phi$ porque $(d(f \circ \Phi))_a = (df)_p \circ (d\Phi)_a = 0 \circ (d\Phi)_a = 0$. En cuanto a la Hessiana, es fácil ver que:

$$\text{Hess}(f \circ \Phi)_a \equiv Q \circ (d\Phi)_a : \mathbb{R}^k \longrightarrow \mathbb{R} ,$$

donde Q es la Hessiana intrínseca de f en el punto crítico p . Como $(d\Phi)_a$ es un isomorfismo, si Q es definida (positiva o negativa) entonces $\text{Hess}(f \circ \Phi)_a$ es definida del mismo signo que Q . Aplicando a $f \circ \Phi$ y a el teorema 74 del apartado 2.9, deducimos que a es un mínimo local estricto para $f \circ \Phi$ si Q es definida positiva o un máximo local estricto si Q es definida negativa. Finalmente, como Φ es biyectiva del entorno W de a al entorno relativo $X \cap E$ de p , deducimos que p es un mínimo (resp. máximo) local estricto de f si a lo es de $f \circ \Phi$. \square

Ejemplo. Sean $X = \{x^2 + y^2 + z^2 = 1\}$ y $f = F|_X$, donde $F(x, y, z) = x^2 + y$. Calculamos:

$$\nabla F = (2x, 1, 0) \quad , \quad \text{Hess}(F) = \begin{bmatrix} 2 & & \\ & 0 & \\ & & 0 \end{bmatrix} \quad , \quad \nabla \frac{x^2 + y^2 + z^2}{2} = (x, y, z) .$$

En cualquier punto $p \in X$, es $T_pX = \left\{ \nabla \left(\frac{x^2 + y^2 + z^2}{2} \right) (p) \right\}^\perp = \{p\}^\perp$.

Las soluciones al problema de multiplicadores de Lagrange:

$$\left. \begin{array}{l} x^2 + y^2 + z^2 = 1 \\ 2x = \lambda x \\ 1 = \lambda y \\ 0 = \lambda z \end{array} \right\}$$

son las siguientes:

$$p^\pm = \left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2}, 0 \right) \quad , \quad \lambda = 2 \quad ,$$

$$q^+ = (0, 1, 0) \quad , \quad \lambda = 1 \quad ,$$

$$q^- = (0, -1, 0) \quad , \quad \lambda = -1 \quad .$$

En los puntos p^\pm , una base de $T_{p^\pm}X$ es $\{v_1, v_2\} = \{(0, 0, 1), (-1, \pm\sqrt{3}, 0)\}$. Como en estos puntos es $\lambda = 2$, para ellos consideramos la matriz:

$$A = \text{Hess}(F) - 2 \cdot \text{Hess} \left(\frac{x^2 + y^2 + z^2}{2} \right) = \begin{bmatrix} 0 & & \\ & -2 & \\ & & -2 \end{bmatrix} ,$$

y entonces la matriz respecto de la base $\{v_1, v_2\}$ de la Hessiana intrínseca en p^\pm es:

$$\begin{bmatrix} v_1^t A v_1 & v_1^t A v_2 \\ v_2^t A v_1 & v_2^t A v_2 \end{bmatrix} = \begin{bmatrix} -2 & \\ & -6 \end{bmatrix} ,$$

definida negativa, luego tanto p^+ como p^- son máximos locales estrictos de $F|_X$.

En los puntos q^\pm , una base de $T_{q^\pm}X$ es $\{w_1, w_2\} = \{(1, 0, 0), (0, 0, 1)\}$. Como $\lambda = 1$ en q^+ , para este punto consideramos la matriz:

$$B^+ = \text{Hess}(F) - 1 \cdot \text{Hess} \left(\frac{x^2 + y^2 + z^2}{2} \right) = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} ,$$

y entonces la matriz respecto de la base $\{w_1, w_2\}$ de la Hessiana intrínseca en q^+ es:

$$\begin{bmatrix} w_1^t B^+ w_1 & w_1^t B^+ w_2 \\ w_2^t B^+ w_1 & w_2^t B^+ w_2 \end{bmatrix} = \begin{bmatrix} 1 & \\ & -1 \end{bmatrix},$$

indefinida, luego q^+ no es ni máximo ni mínimo local de $F|_X$ (es una silla no degenerada, véase el apartado 4.11).

Como en q^- es $\lambda = -1$, para este punto consideramos la matriz:

$$B^- = \text{Hess}(F) - (-1) \cdot \text{Hess}\left(\frac{x^2 + y^2 + z^2}{2}\right) = \begin{bmatrix} 3 & & \\ & 1 & \\ & & 1 \end{bmatrix},$$

y entonces la matriz respecto de la base $\{w_1, w_2\}$ de la Hessiana intrínseca en q^+ es:

$$\begin{bmatrix} w_1^t B^- w_1 & w_1^t B^- w_2 \\ w_2^t B^- w_1 & w_2^t B^- w_2 \end{bmatrix} = \begin{bmatrix} 3 & \\ & 1 \end{bmatrix},$$

definida positiva, luego q^- es un mínimo local estricto de $F|_X$.

Los valores observados:

$$F(p^\pm) = 5/4, \quad F(q^+) = 1, \quad F(p^-) = -1,$$

permiten decir que p^\pm son puntos de máximo, que q^- es punto de mínimo y que p^+ es un punto crítico de valor intermedio. Pero sin las Hessianas intrínsecas no es posible decir si p^\pm son estrictos o no como máximos locales, ni decir si q^- es estricto o no como mínimo local, ni mucho menos decidir si q^+ es máximo local (no blobal) o mínimo local (no global) o ninguna de las dos cosas.

4.11 Lema de Morse

Aquí enunciamos un resultado del matemático estadounidense Marston Morse¹.

Teorema 124. Sean $U \subseteq \mathbb{R}^k$ un abierto y $g : U \rightarrow \mathbb{R}$ de clase C^∞ . Sea $a \in U$ un **punto crítico no degenerado**, es decir que $(dg)_a = 0$ y la Hessiana de g en a es una forma cuadrática Q no degenerada (puede ser indefinida). Entonces en un entorno U^a existe un difeomorfismo $\sigma = (\sigma_1, \dots, \sigma_k)$ con $\sigma(a) = \mathbf{0}$ y tal que $g|_{U^a} = Q(u_1, \dots, u_k)$.

El resultado se pasa de \mathbb{R}^k a funciones en variedades $f : X \rightarrow \mathbb{R}$, sin más que aplicarlo a la compuesta $g = f \circ \Phi$, como hemos hecho en la demostración del teorema 123.

Ejemplo. Supongamos que $a = (a_1, a_2)$ es un punto crítico de $g(x, y)$ y que para todo $v \in \mathbb{R}^2$ es $\text{Hess}(g)_a(v) = v_1^2 - 5v_2^2$, indefinida pero no degenerada. Entonces hay unas coordenadas curvilíneas u_1, u_2 cerca de a tales que $u_1(a) = u_2(a) = 0$ y $g = u_1^2 - 5u_2^2$ en dicho entorno. El grafo de g cerca de a es la imagen de la silla cuadrática $\{z = x^2 - 5y^2\}$ por el siguiente difeomorfismo entre abiertos de \mathbb{R}^3 : $(x, y, z) \mapsto (u_1(x, y), u_2(x, y), z)$, es decir que el grafo de g es una ligera deformación de dicha silla. Las curvas de nivel de g se ven como hipérbolas en las coordenadas u_1, u_2 . Por estas razones decimos que g tiene una **silla no degenerada** en a . Igual ocurre en el punto q^+ del ejemplo del apartado 4.10.

En el entorno de un punto crítico con la hessiana degenerada, la función puede ser inmensamente complicada. Hablar de “sillas degeneradas” no es, pues, del todo adecuado.

¹No confundir con Samuel F. B. Morse, inventor del telégrafo, ni con A. P. Morse, que junto con A. Sard descubrió el teorema de Morse-Sard.