Capítulo 9

Extremos condicionados

Subvariedades diferenciables de $\mathbb{R}^n$. Espacio tangente en un punto. Extremos condicionados: Método de los multiplicadores de Lagrange. Aplicaciones geométricas.

El objetivo central del capítulo es la optimización de funciones reales de varias variables reales, $f(x_1, x_2, \cdots x_n)$, cuando las variables están sometidas a ligaduras:

$$g_1(x_1, x_2, \cdots, x_n) = 0, \quad g_2(x_1, x_2, \cdots, x_n) = 0, \cdots \quad g_m(x_1, x_2, \cdots, x_n) = 0$$

En otras palabras, se trata de calcular los extremos absolutos o relativos (si existen) de f sobre el subconjunto $M \subset \mathbb{R}^n$ formado por los puntos que cumplen las condiciones de ligadura. Cuando este conjunto tiene una estructura geométrica apropiada, que se formula mediante la noción de subvariedad diferenciable, el problema se puede abordar con el método de los multiplicadores de Lagrange.

Por ello el capítulo comienza con el estudio de esta noción geométrica que proporciona el marco natural para los problemas de optimización con restricciones de ligadura en forma de igualdad. Después de estudiar estos problemas se pueden abordar los de optimización con restricciones en forma de desigualdad, de los que no haremos el studio sistemático (con las condiciones de Khun-Tucker) que se puede encontrar en textos más especializados como [10].

Cuando $M \subset \mathbb{R}^n$ sea un subconjunto definido mediante restricciones de desigualdad, el problema de obtener los extremos absolutos de $f|_M$ (que existirán con seguridad cuando f sea continua y M sea compacto) se puede abordar considerando por separado la restricción de M al interior de M, y a su frontera. Lo primero conduce a un problema ordinario de extremos sin ligaduras, que ya han sido considerados en el capítulo 5. La restricción de f a la frontera de M puede conducir a varios problemas de optimización con ligaduras en forma de igualdad: Generalmente la frontera de M no es será subvariedad diferenciable pero habitualmente, cuando M está definido con un número finito de desigualdades, su frontera se puede descomponer en un número finito de subvariedades diferenciables (de diferentes dimensiones). Entonces la optimización de f sobre la frontera de M se podrá atacar con el método de los multiplicadores de Lagrange sobre cada una de las subvariedades diferenciables que componen la frontera.

9.1. Subvariedades diferenciables

En esta sección se introduce la noción de subvariedad diferenciable de $\mathbb{R}^n$, de dimensión k, como un subconjunto $M \subset \mathbb{R}^n$ que localmente tiene una estructura geométrica similar a $\mathbb{R}^k$. Esta estructura geométrica local se puede describir bajo tres formas equivalentes (teorema 9.4) que comenzamos definiendo de manera precisa.

Una de ellas se formula en términos de la gráfica de una función diferenciable de k variables independientes. En las gráficas de este tipo, según la costumbre habitual, las primeras k variables son las independientes que desempeñan un papel especial. Esta restricción artificial se elimina en la siguiente definición considerando cambios de orden en las variables, es decir cambios de variable lineales $T_\sigma : \mathbb{R}^n \to \mathbb{R}^n$,

$$T_\sigma(x_1, x_2, \cdots, x_n) = (x_{\sigma(1)}, x_{\sigma(2)}, \cdots x_{\sigma(n)})$$

asociados a permutaciones $\sigma : \{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\}$.

Definición 9.1 *Diremos que $M \subset \mathbb{R}^n$ admite una representación explícita de clase C^m y dimensión $k < n$, o que M es una gráfica de esa clase y dimensión, si existe una permutación $\sigma : \{1, 2, \cdots, n\} \to \{1, 2, \cdots, n\}$ tal que $M = T_\sigma(G(\mathbf{f}))$ donde*

$$G(\mathbf{f}) = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$$

y $\mathbf{f} : A \to \mathbb{R}^{n-k}$ es de clase C^m en un abierto $A \subset \mathbb{R}^k$.

Otra caracterización de las subvariedades diferenciables de $\mathbb{R}^n$ se expresará en términos de parametrizaciones regulares, que se definen a continuación

Definición 9.2 *Si $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ es una aplicación de clase C^m ($m \geq 1$) definida en un abierto $U \subset \mathbb{R}^k$, con $1 \leq k \leq n$, se dice que $\boldsymbol{\varphi}$ es una parametrización de clase C^m y dimensión k. Si además se cumplen las condiciones*

i) $\boldsymbol{\varphi}$ es un homeomorfismo entre U y su imagen $\boldsymbol{\varphi}(U)$ (con la topología relativa).

ii) Para cada $\mathbf{u} \in U$, los vectores $D_1\boldsymbol{\varphi}(\mathbf{u}), D_2\boldsymbol{\varphi}(\mathbf{u}), \cdots, D_k\boldsymbol{\varphi}(\mathbf{u})$ son linealmente independientes.

se dice que $\boldsymbol{\varphi}$ es una parametrización regular. En este caso, si $M = \boldsymbol{\varphi}(U)$ diremos que $\boldsymbol{\varphi}$ es una parametrización regular de M y también que $M \subset \mathbb{R}^n$ admite una parametrización regular (de clase C^m y dimensión $k \leq n$).

Toda representación explícita (de clase C^m y dimensión $k < n$) de $M \subset \mathbb{R}^n$ lleva asociada de modo natural una parametrización regular de M de la misma clase y dimensión, $\boldsymbol{\varphi}(\mathbf{u}) = T_\sigma((\mathbf{u}, \mathbf{f}(\mathbf{u})))$, definida en $U = A$, es decir:

$$\boldsymbol{\varphi}(u_1, u_2, \cdots u_k) = T_\sigma(u_1, u_2, \cdots, u_k, f_1(u_1, \cdots, u_k), \cdots f_{n-k}(u_1, \cdots, u_k))$$

Dejamos como ejercicio al cuidado del lector la comprobación de que se cumplen las condiciones requeridas en la definición 9.2.

Un ejemplo sencillo de parametrización regular lo proporciona la parametrización habitual de un trozo de esfera mediante la longitud y la latitud. Esto se puede ver en H.7 donde también se muestra un ejemplo interesante de lo que puede ocurrir cuando una parametrización no es regular aunque sea inyectiva.

Definición 9.3 *Diremos que $M \subset \mathbb{R}^n$ admite una representación implícita de clase C^m y dimensión $k < n$ si se puede representar en la forma*

$$M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$$

donde $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$ es de clase C^m en un abierto $\Omega \subset \mathbb{R}^n$, y para cada $\mathbf{x} \in M$ los vectores $\nabla g_1(\mathbf{x})$, $\nabla g_2(\mathbf{x})$, $\cdots$, $\nabla g_{n-k}(\mathbf{x})$ son linealmente independientes (e.d. la matriz $(D_i g_j(\mathbf{x}))_{1\leq i\leq n, 1\leq j\leq n-k}$ tiene rango $n-k$).

Toda representación explícita de clase C^m y dimensión $k < n$ de $M \subset \mathbb{R}^n$ proporciona de modo natural una representación implícita de la misma clase y dimensión, dada en términos de la función:

$$\mathbf{g}(\mathbf{u},\mathbf{v}) = \mathbf{f}(\mathbf{u}) - \mathbf{v} = (f_1(u_1, u_2, \cdots u_k) - v_1, \cdots, f_{n-k}(u_1, u_2, \cdots u_k) - v_{n-k})$$

definida en $\Omega = A \times \mathbb{R}^{n-k}$. Dejamos al cuidado del lector la comprobación de que se cumplen las condiciones requeridas en la definición 9.3.

Teorema 9.4 *Si $1 \leq k < n$, las siguientes propiedades de $M \subset \mathbb{R}^n$ son equivalentes*

a) Cada $\mathbf{p} \in M$ posee un entorno abierto $\Omega_\mathbf{p} \subset \mathbb{R}^n$ tal que $M \cap \Omega_\mathbf{p}$ admite una parametrización regular de clase C^m y dimensión k.

b) Cada $\mathbf{p} \in M$ posee un entorno abierto $\Omega_\mathbf{p} \subset \mathbb{R}^n$ tal que $M \cap \Omega_\mathbf{p}$ admite una representación implícita clase C^m y dimensión k.

c) Cada $\mathbf{p} \in M$ posee un entorno abierto $\Omega_\mathbf{p} \subset \mathbb{R}^n$ tal que $M \cap \Omega_\mathbf{p}$ admite una representación explícita de clase C^m y dimensión k.

Dem: Según lo comentado después de las definiciones 9.2 y 9.3 es claro que c) $\Rightarrow$ a), y c) $\Rightarrow$ b).
b) $\Rightarrow$ c) es consecuencia directa del teorema de la función implícita: Por hipótesis, cada $\mathbf{p} \in M$ posee un entorno abierto $\Omega_\mathbf{p} \subset \mathbb{R}^n$ tal que

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

donde $\mathbf{g}$ está definida en $\Omega_\mathbf{p}$ y cumple las condiciones de la definición 9.3. Como la matriz

$$\begin{pmatrix} D_1 g_1(\mathbf{p}) & D_2 g_1(\mathbf{p}) & \cdots & D_n g_1(\mathbf{p}) \\ \cdots & \cdots & \cdots & \cdots \\ D_1 g_{n-k}(\mathbf{p}) & D_2 g_{n-k}(\mathbf{p}) & \cdots & D_n g_{n-k}(\mathbf{p}) \end{pmatrix}$$

tiene rango $n-k$ podemos suponer, para simplificar la notación, que no es nulo el determinante de la matriz cuadrada formada con las últimas $n-k$ columnas. Entonces, aplicando el teorema de la función implícita, se puede asegurar que existe un entorno abierto de $\mathbf{p}$, de la forma $\Omega'_\mathbf{p} = A \times B \subset \Omega_\mathbf{p}$, con $A \subset \mathbb{R}^k$, $B \subset \mathbb{R}^{n-k}$, abiertos, y una función implícita de clase C^m, $\mathbf{f} : A \to B$ tal que

$$M \cap \Omega'_\mathbf{p} = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$$

Si fuese nulo el determinante formado por las últimas $n-k$ columnas, tendríamos que seleccionar otras columnas para conseguir una matriz cuadrada de determinante no nulo y el teorema de la función implícita permitiría despejar localmente las correspondientes $n-k$ variables en función de las restantes. Con una permutación σ de $\{1,2,\cdots,n\}$ se consigue que estas $n-k$ variables pasen a ser las últimas, y esta permutación es la que hace que se cumpla la definición 9.1.
a) $\Rightarrow$ c) Se demuestra con ayuda del teorema de la función inversa: Por hipótesis $\mathbf{p} \in M$ posee un entorno abierto $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$ tal que $M \cap \Omega_{\mathbf{p}} = \boldsymbol{\varphi}(U)$, donde $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ es una parametrización regular de clase C^m y dimensión k definida en un abierto $U \subset \mathbb{R}^k$. Si $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$, como los vectores $D_1\boldsymbol{\varphi}(\mathbf{a})$, $D_2\boldsymbol{\varphi}(\mathbf{a})$, $\cdots$, $D_k\boldsymbol{\varphi}(\mathbf{a})$ son linealmente independientes, la matriz

$$\begin{pmatrix} D_1\varphi_1(\mathbf{a}) & D_2\varphi_1(\mathbf{a}) & \cdots & D_k\varphi_1(\mathbf{a}) \\ \cdots & \cdots & \cdots & \cdots \\ D_1\varphi_n(\mathbf{a}) & D_2\varphi_n(\mathbf{a}) & \cdots & D_k\varphi_n(\mathbf{a}) \end{pmatrix}$$

tiene rango k, y podemos suponer, para simplificar la escritura, que la matriz formada por la k primeras filas tiene determinante no nulo. En ese caso escribimos

$$\boldsymbol{\varphi}(\mathbf{u}) = (\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u}))$$

donde $\boldsymbol{\alpha} : U \to \mathbb{R}^k$, $\boldsymbol{\beta} : U \to \mathbb{R}^{n-k}$.

Como $\det \alpha'(\mathbf{a}) \neq 0$, aplicando el teorema de la función inversa a la aplicación $\boldsymbol{\alpha}$ en el punto $\mathbf{a}$ obtenemos $A \subset U$, entorno abierto de $\mathbf{a}$, tal que $\boldsymbol{\alpha}(A) \subset \mathbb{R}^k$ es abierto y $\boldsymbol{\alpha} : A \to \boldsymbol{\alpha}(A)$ es de clase C^m con inversa de clase C^m. Entonces $\mathbf{f} = \boldsymbol{\beta} \circ \boldsymbol{\alpha}^{-1}$, definida en $\boldsymbol{\alpha}(A)$, con valores en $\mathbb{R}^{n-k}$, también es de clase C^m.

Como $\boldsymbol{\varphi}$ establece un homeomorfismo entre U y $\boldsymbol{\varphi}(U)$ se tiene que $\boldsymbol{\varphi}(A)$ es un entorno abierto del punto $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ en el conjunto $\boldsymbol{\varphi}(U) = M \cap \Omega_{\mathbf{p}}$, para la topología relativa de este conjunto, luego es de la forma $\boldsymbol{\varphi}(A) = M \cap \Omega'_{\mathbf{p}}$, donde $\Omega'_{\mathbf{p}} \subset \mathbb{R}^n$ es un abierto que se puede suponer incluido en $\Omega_{\mathbf{p}}$. Este abierto cumple las condiciones requeridas en la definición 9.1:

$$M \cap \Omega'_{\mathbf{p}} = \boldsymbol{\varphi}(A) = \{(\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u})) : \mathbf{u} \in A\} = \{(\mathbf{v}, \mathbf{f}(\mathbf{v})) : \mathbf{v} \in \boldsymbol{\alpha}(A)\}$$

En caso de que fuese nulo el determinante de la matriz formada con las k primeras filas de la matriz $(D_i\boldsymbol{\varphi}_j(\mathbf{a}))$ tendríamos que seleccionar otras filas para obtener una matriz cuadrada de determinante no nulo. Estas filas se pueden llevar a las primeras posiciones mediante una permutación σ de las variables, y con esta permutación se cumplen las condiciones de la definición 9.1. ■

Observación: Sea $M = \boldsymbol{\varphi}(U)$, donde $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ es una parametrización regular de clase C^m y dimensión k. Según la definición 9.2 la inversa $\boldsymbol{\varphi}^{-1} : M \to U$ es continua. Ahora podemos decir algo más: Cada $\mathbf{p} \in M$ posee un entorno abierto $W_{\mathbf{p}} \subset \mathbb{R}^n$, donde hay definida una función de clase C^m, $\boldsymbol{\Psi} : W_{\mathbf{p}} \to U$, que verifica

$$\boldsymbol{\Psi}(\mathbf{x}) = \boldsymbol{\varphi}^{-1}(\mathbf{x}) \quad \text{para todo } \mathbf{x} \in M \cap W_{\mathbf{p}}.$$

Para ver esto basta continuar con el razonamiento de a) $\Rightarrow$ c) en el teorema anterior y definir en $W_{\mathbf{p}} = \boldsymbol{\alpha}(A) \times \mathbb{R}^{n-k}$ la aplicación $\boldsymbol{\Psi}(\mathbf{s}, \mathbf{t}) = \boldsymbol{\alpha}^{-1}(\mathbf{s})$. Es claro que para cada $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{u}) = (\boldsymbol{\alpha}(\mathbf{u}), \boldsymbol{\beta}(\mathbf{u})) \in W_{\mathbf{p}}$ se cumple $\boldsymbol{\varphi}^{-1}(\mathbf{x}) = \mathbf{u} = \boldsymbol{\Psi}(\mathbf{x})$.

Definición 9.5 *Un subconjunto $M \subset \mathbb{R}^n$ se dice que es una subvariedad diferenciable de clase C^m y dimensión k, ($1 \leq k < n$), cuando posee las propiedades equivalentes del teorema 9.4*

Ejemplos 9.6 *En cada uno de los siguientes casos $M \subset \mathbb{R}^n$ es una subvariedad de clase C^m y dimensión k:*

a) *$M = \boldsymbol{\varphi}(U)$ donde $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ es una parametrización regular de clase C^m definida en un abierto $U \subset \mathbb{R}^k$.*

b) *$M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$ donde $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$ es una aplicación de clase C^m, definida en un abierto $\Omega \subset \mathbb{R}^n$, con la propiedad de que, para cada $\mathbf{x} \in M$ los vectores $\nabla g_1(\mathbf{x}), \nabla g_2(\mathbf{x}), \cdots, \nabla g_{n-k}(\mathbf{x})$ son linealmente independientes.*

c) *$M = T_\sigma(G(\mathbf{f}))$ donde $G(\mathbf{f}) = \{(\mathbf{u}, \mathbf{f}(\mathbf{u})) : \mathbf{u} \in A\}$ es la gráfica de una aplicación de clase C^m, $\mathbf{f} : A \to \mathbb{R}^{n-k}$, definida en un abierto $A \subset \mathbb{R}^k$, y σ es una permutación de $\{1, 2, \cdots, n\}$.*

Espacio tangente a una subvariedad diferenciable. Según la definición 5.3 un vector $\mathbf{w} \in \mathbb{R}^n$ es tangente a $M \subset \mathbb{R}^n$ en $\mathbf{p} \in M$ si existe una aplicación $\boldsymbol{\gamma} : (-\alpha, \alpha) \to M$, derivable en $t = 0$ tal que $\boldsymbol{\gamma}(0) = \mathbf{p}$ y $\boldsymbol{\gamma}'(0) = \mathbf{w}$. Si $T_\mathbf{p}(M)$ es el conjunto de los vectores tangentes a M en $\mathbf{p}$ y $\Omega_\mathbf{p}$ es un entorno abierto de $\mathbf{p}$, es inmediato que $T_\mathbf{p}(M) = T_\mathbf{p}(M \cap \Omega_\mathbf{p})$.

Si $A \subset \mathbb{R}^k$ es abierto y $\mathbf{f} : A \to \mathbb{R}^{n-k}$ es diferenciable en $\mathbf{a} \in A$, según la proposición 5.28, el conjunto de vectores tangentes en $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$ a la gráfica

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

es la gráfica de la diferencial $d\mathbf{f}(\mathbf{a})$, es decir el subespacio vectorial, de dimensión k

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

Cuando $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$ con $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$ definida en un abierto $\Omega \subset \mathbb{R}^n$, y diferenciable en $\mathbf{p} \in M$, según la proposición 5.29 se verifica

$$T_\mathbf{p}(M) \subset \mathrm{Ker}\ d\mathbf{g}(\mathbf{p}) = \bigcap_{j=1}^{n-k} \{\mathbf{u} \in \mathbb{R}^n : \langle\ \nabla g_j(\mathbf{p}) \mid \mathbf{u}\ \rangle = 0\}$$

Por consiguiente, cuando los vectores $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots, \nabla g_{n-k}(\mathbf{p})$ son linealmente independientes Ker $d\mathbf{g}(\mathbf{p})$ es un subespacio de dimensión k que contiene a $T_\mathbf{p}(M)$.

Proposición 9.7 *Sea $M = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = 0, \cdots g_{n-k}(\mathbf{x}) = 0\}$, donde las funciones $g_1, g_2 \cdots g_{n-k} : \Omega \to \mathbb{R}$ son de clase $C^1(\Omega)$. Si $\mathbf{p} \in M$ y los vectores*

$$\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots \nabla g_{n-k}(\mathbf{p})$$

son linealmente independientes entonces

$$T_\mathbf{p}(M) = \bigcap_{j=1}^{n-k} \{\mathbf{u} \in \mathbb{R}^n :\ \langle\ \nabla g_j(\mathbf{p}) \mid \mathbf{u}\ \rangle = 0\}$$

Es decir, $T_\mathbf{p}(M)$ es el subespacio vectorial de $\mathbb{R}^n$, de dimensión k, ortogonal a los vectores $\nabla g_1(\mathbf{p})$, $\nabla g_2(\mathbf{p})$.... $\nabla g_{n-k}(\mathbf{p})$.

Dem: Como $T_{\mathbf{p}}(M) \subset \operatorname{Ker} d\mathbf{g}(\mathbf{p})$, y $\operatorname{Ker} d\mathbf{g}(\mathbf{p})$ es un subespacio vectorial de dimensión k, para obtener la igualdad $T_{\mathbf{p}}(M) = \operatorname{Ker} d\mathbf{g}(\mathbf{p})$ basta ver $T_{\mathbf{p}}(M)$ es un subespacio vectorial de la misma dimensión.

Razonando como en la demostración de b) $\Rightarrow$ c) en el teorema 9.4, el teorema de la función implícita permite obtener $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$, entorno abierto de $\mathbf{p}$, con tal que $M \cap \Omega_{\mathbf{p}}$, es (salvo una permutación de las variables) la gráfica de una aplicación $\mathbf{f} : A \to \mathbb{R}^{n-k}$, diferenciable en $\mathbf{a} \in A$, donde $\mathbf{p} = (\mathbf{a}, \mathbf{f}(\mathbf{a}))$. Basta tener en cuenta ahora que $T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$ es un espacio vectorial de dimensión k y que $T_{\mathbf{p}}(M) = T_{\mathbf{p}}(\Omega_{\mathbf{p}} \cap M)$. ■

Finalmente, consideramos el caso de un conjunto de la forma $M = \boldsymbol{\varphi}(U)$, donde $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ está definida en un abierto $U \subset \mathbb{R}^k$ y es diferenciable en $\mathbf{a} \in U$. Recordemos que el conjunto de vectores tangentes a M en $\mathbf{p} = \boldsymbol{\varphi}(\mathbf{a})$ contiene al subespacio $d\boldsymbol{\varphi}(\mathbf{a})(\mathbb{R})^k$, es decir, $T_{\mathbf{p}}(M) \supset \{D_{\mathbf{u}}\boldsymbol{\varphi}(\mathbf{a}) : \mathbf{u} \in \mathbb{R}^k\}$. La dimensión de este subespacio es $\leq k$ y será igual a k, cuando las derivadas parciales $D_j\boldsymbol{\varphi}(\mathbf{a}) = d\boldsymbol{\varphi}(\mathbf{a})\mathbf{e}_j$, $1 \leq j \leq k$, sean linealmente independientes. En este caso, si se sabe que $T_{\mathbf{p}}(M)$ es un subespacio vectorial de dimensión k se obtendrá que

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

Esto es lo que ocurre cuando $\boldsymbol{\varphi}$ es una parametrización regular de clase C^m y dimensión k.

Proposición 9.8 *Si $M \subset \mathbb{R}^n$ es una subvariedad diferenciable de clase C^m y dimensión k entonces en cada punto $\mathbf{p} \in M$ el espacio tangente $T_{\mathbf{p}}(M)$ es un espacio vectorial de dimensión k.*

Dem: Basta considerar que M tiene localmente la propiedad 9.4 c): Existe $\Omega_{\mathbf{p}} \subset \mathbb{R}^n$, entorno abierto de $\mathbf{p}$, tal que $M \cap \Omega_{\mathbf{p}} = \{(\mathbf{x}, \mathbf{f}(\mathbf{x}) : \mathbf{x} \in A\}$ donde $\mathbf{f} : A \to \mathbb{R}^{n-k}$ es diferenciable en el abierto $A \subset \mathbb{R}^{n-k}$. Se sigue de esto que

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

luego $T_{\mathbf{p}}(M)$ es un espacio vectorial de dimensión k. (Para simplificar la notación, al aplicar la condición 9.4 c) hemos supuesto σ = identidad) ■

Cuando $M \subset \mathbb{R}^n$ es una subvariedad diferenciable de clase C^m y dimensión k, podemos aplicar en cada punto $\mathbf{p} \in M$ lo indicado anteriormente. Así, para escribir les ecuaciones en forma implícita del espacio tangente $T_{\mathbf{p}}(M)$ podemos utilizar, en un entorno abierto $\Omega_{\mathbf{p}}$ de $\mathbf{p}$, una representación implícita de $M \cap \Omega_{\mathbf{p}}$. Análogamente, para escribir las ecuaciones paramétricas de $T_{\mathbf{p}}(M)$, podemos utilizar una parametrización regular de un entorno relativo de $\mathbf{p}$ (véase la sección 5.5).

9.2. Extremos condicionados

Dada una función $f : \Omega \to \mathbb{R}$ definida en un abierto $\Omega \subset \mathbb{R}^n$ se trata de determinar, si existen, los extremos (relativos o absolutos) de la restricción de f a un conjunto $M \subset \Omega$ de la forma $M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = 0\}$ donde $\mathbf{g} : \Omega \to \mathbb{R}^m$. Es decir, se trata de obtener los extremos (relativos o absolutos) de $f(x_1, x_2, \cdots x_n)$ cuando las variables $(x_1, x_2, \cdots x_n)$ están sometidas a las condiciones de ligadura

$$g_1(x_1, x_2, \cdots x_n) = 0, \quad g_2(x_1, x_2, \cdots x_n) = 0, \cdots \ g_m(x_1, x_2, \cdots x_n) = 0.$$

Las definiciones referentes a extremos condicionados son obvias: Si $\mathbf{p} \in M$ posee un entorno abierto $V_{\mathbf{p}} \subset \Omega$ tal que $f(\mathbf{x}) \leq f(\mathbf{p})$ (resp. $f(\mathbf{x}) \geq f(\mathbf{p})$ para todo $\mathbf{x} \in M \cap V_{\mathbf{p}}$, se dice que $f|_M$ presenta en $\mathbf{p}$ un máximo (resp. mínimo) relativo o bien que f presenta en $\mathbf{p}$ un máximo (resp. mínimo) condicionado por las ligaduras $g_k(x_1, x_2, \cdots x_n) = 0$, $1 \leq k \leq m$. Cuando se reemplaza el signo de desigualdad $\leq$ por el de desigualdad estricta $<$, se obtiene la definición de máximo (resp. mínimo) relativo estricto condicionado.

Los principales resultados sobre extremos condicionados se obtienen cuando f es diferenciable y el conjunto M definido por las condiciones de ligadura, tiene una estructura geométrica sencilla desde el punto de vista del cálculo diferencial: Las propiedades que se exigirán a las condiciones de ligadura g_k, $1 \leq k \leq m$, serán las adecuadas para garantizar que en cada $\mathbf{p} \in M$ el conjunto $T_{\mathbf{p}}(M)$ es un espacio vectorial de dimensión $n - m$ (véase 9.7).

Para motivar los resultados generales consideramos primero el caso particular para $n = 2$, $m = 1$. En este caso $M = \{(x, y) \in \Omega : g(x, y) = 0\}$ es una "curva" definida implícitamente por una función g de clase C^1 en un abierto $\Omega \subset \mathbb{R}^2$.

Si $\mathbf{p} = (a, b) \in M$ es un punto donde $\nabla g(\mathbf{p}) = (D_1g(a, b), D_2g(a, b)) \neq (0, 0)$, en virtud de la proposición 9.7 el espacio tangente $T_{\mathbf{p}}(M)$ es el subespacio de $\mathbb{R}^2$, de dimensión 1, ortogonal al gradiente $\nabla g(\mathbf{p})$. Es decir, existe la recta tangente a la "curva" M en el punto $\mathbf{p}$, y su ecuación es

$$D_1g(a, b)(x - a) + D_2g(a, b)(y - b) = 0$$

Representemos la función $f : \Omega \to \mathbb{R}$ dibujando en el plano sus curvas de nivel $N_c = \{(x, y) \in \Omega : f(x, y) = c\}$, para diferentes valores de c. Si $\mathbf{p} \in \Omega$ y $c = f(\mathbf{p})$ entonces N_c es la curva de nivel que pasa por $\mathbf{p} \in \Omega$ y sabemos que $\nabla f(\mathbf{p})$ es ortogonal todos los vectores tangentes a N_c en $\mathbf{p}$. Supongamos además que existe un entorno $B(\mathbf{p}, r) \subset \Omega$ que queda descompuesto por la curva N_c en dos regiones separadas:

$$\{(x, y) \in B(\mathbf{p}, r) : f(x, y) < c\}, \quad \{(x, y) \in B(\mathbf{p}, r) : f(x, y) > c\}$$

que llamaremos, respectivamente, lado izquierdo y lado derecho de N_c. Si la curva M pasa por $\mathbf{p}$ atravesando la curva de nivel N_c entonces $f(x, y) - c$ cambia de signo cuando $(x, y) \in M$ pasa de un lado al otro de N_c, luego $f|_M$ no puede presentar en $\mathbf{p}$ ni un máximo ni un mínimo relativo. Por lo tanto, para que $f|_M$ presente en $\mathbf{p} \in M$ un extremo relativo, es necesario que la curva M no atraviese a la línea de nivel N_c, es decir M y N_c deben tener en $\mathbf{p}$ la misma recta tangente. Esto ocurre si y sólo sí $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$ para algún $\mu \in \mathbb{R}$. Lo que acabamos de ver es una justificación heurística, y una interpretación geométrica, de la siguiente proposición

Proposición 9.9 *Sea $M = \{(x, y) \in \Omega : g(x, y) = 0\}$, donde $g : \Omega \to \mathbb{R}$ es de clase $C^1(\Omega)$ en un abierto $\Omega \subset \mathbb{R}^2$. Se supone que $f : \Omega \to \mathbb{R}$ es diferenciable en $\mathbf{p} \in M$ y que $\nabla g(\mathbf{p}) \neq (0, 0)$. Si $f|_M$ presenta un extremo relativo en $\mathbf{p}$ entonces existe $\mu \in \mathbb{R}$ tal que $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$.*

DEM: Según la proposición 9.7 el subespacio ortogonal a $T_{\mathbf{p}}(M)$ es el generado por $\nabla g(\mathbf{p}) \neq (0, 0)$. Por lo tanto basta demostrar que si $f|_M$ presenta en $\mathbf{p} \in M$ un extremo relativo entonces $\nabla f(\mathbf{p})$ es ortogonal a $T_{\mathbf{p}}(M)$.

Supongamos que el extremo relativo es un máximo y sea $V_{\mathbf{p}} \subset \Omega$ un entorno abierto de $\mathbf{p}$ tal que $f(\mathbf{x}) \leq f(\mathbf{p})$ para todo $\mathbf{x} \in V_{\mathbf{p}}$.

Dado $\mathbf{u} \in T_{\mathbf{p}}(M)$, según la definición de vector tangente, existe una aplicación $\boldsymbol{\gamma} : (-\delta, \delta) \to M$ derivable en $t = 0$, tal que $\boldsymbol{\gamma}(0) = \mathbf{p}$ y $\boldsymbol{\gamma}'(0) = \mathbf{u}$. Como $\boldsymbol{\gamma}$ es continua en $t = 0$ existe $0 < r < \delta$ tal que $|t| < r \Rightarrow \boldsymbol{\gamma}(t) \in V_{\mathbf{p}} \cap M$, luego

$$|t| < r \Rightarrow f(\boldsymbol{\gamma}(t)) \leq f(\boldsymbol{\gamma}(0))$$

La función real de variable real $\varphi(t) = f(\boldsymbol{\gamma}(t))$ presenta un máximo relativo en $t = 0$ donde es derivable (en virtud de la regla de la cadena) luego

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

Es decir, $\nabla f(\mathbf{p})$ es ortogonal a $T_{\mathbf{p}}(M)$. ■

En las condiciones de la proposición anterior, los puntos $\mathbf{p} \in M$ que cumplen $\nabla g(\mathbf{p}) \neq (0, 0)$ diremos que son puntos regulares de M. Los puntos regulares $\mathbf{p}$ donde f es diferenciable y $\nabla f(\mathbf{p}) = \mu \nabla g(\mathbf{p})$ para algún $\mu \in \mathbb{R}$ se dice que son *puntos estacionarios* de $f|_M$. Denotaremos por $E(f, M)$ el conjunto de todos los puntos estacionarios de $f|_M$. Si $\mathbf{p} \in E(f, M)$, el coeficiente μ que hace que se cumplan las igualdades

$$D_1 f(\mathbf{p}) = \mu D_1 g(\mathbf{p}), \quad D_2 f(\mathbf{p}) = \mu D_2 g(\mathbf{p})$$

se llama multiplicador de Lagrange asociado al punto estacionario $\mathbf{p}$. Cada punto estacionario $\mathbf{p} \in E(f, M)$ lleva asociado su multiplicador de Lagrange.

Con esta terminología la proposición 9.9 se enuncia diciendo que los puntos regulares $\mathbf{p} \in M$ donde una función diferenciable f presenta extremos relativos condicionados por M son puntos estacionarios de $f|_M$.

En las aplicaciones habituales f es diferenciable en todos los puntos de M y todos los puntos de M son regulares. En este caso los extremos relativos de $f|_M$ se alcanzan en puntos estacionario y para determinar los extremos relativos de $f|_M$ se comienza calculando el conjunto de los puntos estacionarios $E(f, M)$. Esto se consigue resolviendo (cuando sea posible) el sistema de tres ecuaciones con tres incógnitas, x, y, μ.

$$D_1 f(x, y) - \mu D_1 g(x, y) = 0; \quad D_2 f(x, y) - \mu D_2 g(x, y) = 0; \quad g(x, y) = 0.$$

Si el sistema no tiene solución se puede asegurar que $f|_M$ no presenta extremos relativos. Si el sistema tiene un número finito de soluciones

$$(x_1, y_1, \mu_1), (x_2, y_2, \mu_2), \cdots (x_k, y_k, \mu_k)$$

tendremos un conjunto finito de puntos estacionarios

$$\mathbf{p}_1 = (x_1, y_1), \quad \mathbf{p}_2 = (x_2, y_2), \ \cdots \mathbf{p}_k = (x_k, y_k)$$

y sus correspondientes multiplicadores $\mu_1, \ \mu_2, \ \cdots \mu_k$.

Si además M es compacto y el problema que nos ocupa consiste en determinar el máximo y el mínimo absoluto de f sobre M, bastará evaluar f en cada uno de los puntos estacionarios para obtener, por inspección directa, los extremos absolutos:

$$\max_{\mathbf{x}\in M} f(\mathbf{x}) = \max\{f(\mathbf{p}_i) : 1 \leq i \leq k\}$$

$$\min_{\mathbf{x}\in M} f(\mathbf{x}) = \min\{f(\mathbf{p}_i) : 1 \leq i \leq k\}$$

Lo mismo se podrá hacer cuando M no sea compacto pero por la naturaleza del problema sepamos que f debe alcanzar en M un máximo o un mínimo absoluto, o ambos. Así por ejemplo, como M es cerrado, la función $f(x,y) = (x-a)^2 + (y-b)^2$ siempre alcanza un mínimo absoluto sobre M (recuérdese que, según lo indicado a continuación de proposición 3.17, fijado un punto $\mathbf{a} \in \mathbb{R}^n$, la función distancia $\mathbf{x} \to \|\mathbf{x} - \mathbf{a}\|_2$ siempre alcanza un mínimo absoluto sobre cada cerrado $M \subset \mathbb{R}^n$).

A la hora de las aplicaciones concretas, cuando f no sea diferenciable en todos los puntos de M o algún punto de M no sea regular, para determinar los extremos absolutos de f se puede proceder como antes, pero considerando también los puntos de M donde f no es diferenciable o g no es regular. Si el conjunto de estos puntos $S(f, M)$ es finito y sabemos que existe un máximo, o un mínimo absolutos, lo podremos calcular como antes evaluando la función en un conjunto finito de puntos:

$$\max_{\mathbf{x}\in M} f(\mathbf{x}) = \max\{f(\mathbf{x}) : \mathbf{x} \in E(f, M) \cup S(f, M)\}$$

$$\min_{\mathbf{x}\in M} f(\mathbf{x}) = \min\{f(\mathbf{x}) : \mathbf{x} \in E(f, M) \cup S(f, M)\}$$

Un problema más delicado es el de determinar, entre los puntos estacionarios los que proporcionan un máximo relativo, o un mínimo relativo. Más adelante daremos condiciones de segundo orden suficientes para que un punto estacionario sea un punto de máximo relativo o un punto de mínimo relativo. Pero este criterio no será aplicable a los puntos donde f o g no sean dos veces diferenciables.

En el caso que nos ocupa, para funciones de dos variables reales, dado un punto estacionario $\mathbf{p} \in M$ se puede proceder de la siguiente forma: En un entorno $V_\mathbf{p}$ de $\mathbf{p}$ se busca una parametrización del trozo de curva $V_\mathbf{p} \cap M$, es decir un homeomorfismo diferenciable $\boldsymbol{\varphi} : (\alpha, \beta) \to M \cap V_\mathbf{p}$. Si $\boldsymbol{\varphi}(t_0) = \mathbf{p}$ es claro que $\mathbf{p}$ será un punto de máximo (resp. mínimo) relativo para $f|_M$ si y sólo sí t_0 es un punto de máximo (resp. mínimo) de la función real de variable real $f(\boldsymbol{\varphi}(t))$. Este último estudio se podrá abordar con las técnicas usuales de las funciones de una variable.

La parametrización local $\boldsymbol{\varphi}$ a veces se puede conseguir considerando la ecuación de la curva $g(x,y) = 0$ y despejando, en un entorno de $\mathbf{p}$ suficientemente pequeño, una variable en función de la otra, es decir obteniendo un entorno rectangular de $\mathbf{p}$, $V_{\mathbf{p}} = (a,b) \times (c,d)$, tal que $V_{\mathbf{p}} \cap M$ se pueda expresar de una de las dos formas siguientes

$$\{(x,\alpha(x)) : a < x < b\}; \quad \{(\beta(y), y) : c < y < d\}$$

(El teorema de la función implícita nos asegura que esto siempre se puede hacer) En el primer caso se consigue la parametrización local $\boldsymbol{\varphi}(t) = (t, \alpha(t))$, $a < t < b$ y en el segundo caso $\boldsymbol{\varphi}(t) = (\beta(t), t)$, $c < t < d$.

Después de haber visto con detalle el caso particularmente simple de las funciones de dos variables sometidas a una ligadura consideramos el caso general de los extremos relativos de una función de n variables sometida a m condiciones de ligadura ($m < n$).

Teorema 9.10 *Sea $\Omega \subset \mathbb{R}^n$ abierto y $M = \{\mathbf{x} \in \Omega : g_k(\mathbf{x}) = 0,\ 1 \leq k \leq m\}$ donde las funciones $g_k : \Omega \to \mathbb{R}$ son de clase $C^1(\Omega)$. Se supone que $f : \Omega \to \mathbb{R}$ es diferenciable en $\mathbf{p} \in M$ y que los vectores $\nabla g_1(\mathbf{p})$, $\nabla g_2(\mathbf{p})$,... $\nabla g_m(\mathbf{p})$ son linealmente independientes.*

Si $f|_M$ presenta en $\mathbf{p} \in M$ un extremo relativo entonces $df(\mathbf{p})$ se anula sobre el espacio tangente $T_{\mathbf{p}}(M)$, y por lo tanto existen coeficientes $\mu_1, \mu_2, \cdots, \mu_m \in \mathbb{R}$ tales que

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \cdots + \mu_m \nabla g_m(\mathbf{p})$$

Dem: Supongamos que $f|_M$ presenta en $\mathbf{p}$ un máximo relativo y sea $V_{\mathbf{p}} \subset \Omega$ un entorno abierto de $\mathbf{p}$ tal que $f(\mathbf{x}) \leq f(\mathbf{p})$ para todo $\mathbf{x} \in V_{\mathbf{p}}$.

Dado $\mathbf{u} \in T_{\mathbf{p}}(M)$, según la definición de vector tangente, existe una aplicación $\boldsymbol{\gamma} : (-\delta, \delta) \to M$ derivable en $t = 0$, tal que $\boldsymbol{\gamma}(0) = \mathbf{p}$, y $\boldsymbol{\gamma}'(0) = \mathbf{u}$. Como $\boldsymbol{\gamma}$ es continua en $t = 0$ existe $0 < r < \delta$ tal que $|t| < r \Rightarrow \boldsymbol{\gamma}(t) \in V_{\mathbf{p}} \cap M$, luego

$$|t| < r \Rightarrow f(\boldsymbol{\gamma}(t)) \leq f(\boldsymbol{\gamma}(0))$$

La función real de variable real $\varphi(t) = f(\boldsymbol{\gamma}(t))$ presenta un máximo relativo en $t = 0$ donde (en virtud de la regla de la cadena) es derivable luego

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

Si $E \subset \mathbb{R}^n$ es el subespacio m-dimensional generado por los vectores linealmente independientes $\nabla g_k(\mathbf{p})$, $1 \leq k \leq m$, la proposición 9.7 asegura que $T_{\mathbf{p}}(M)$ es el subespacio ortogonal a E. Hemos demostrado que $\nabla f(\mathbf{p})$ es ortogonal a $T_{\mathbf{p}}(M)$, luego $\nabla f(\mathbf{p}) \in E$ se puede expresar como una combinación lineal de los vectores $\nabla g_k(\mathbf{p})$, $1 \leq k \leq m$, que forman una base de E. ∎

De forma similar al caso $n = 2$, $m = 1$, los puntos $\mathbf{p} \in M$ donde los vectores $\nabla g_1(\mathbf{p})$, $\nabla g_2(\mathbf{p})$,..$\nabla g_m(\mathbf{p})$, son linealmente independientes diremos que son *puntos regulares* de $\mathbf{g}$. Los puntos regulares $\mathbf{p}$ donde f es diferenciable y $\nabla f(\mathbf{p})$ es combinación lineal de estos vectores se dice que son *puntos estacionarios* de $f|_M$.

Denotaremos por $E(f,M)$ el conjunto de tales puntos. Para cada $\mathbf{p} \in E(f,M)$, los coeficientes μ_k, $1 \leq k \leq m$, que hacen que se cumpla la igualdad

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \cdots + \mu_m \nabla g_m(\mathbf{p})$$

se llaman *multiplicadores de Lagrange* asociados al punto estacionario $\mathbf{p}$. Cada punto estacionario $\mathbf{p} \in E(f,M)$ lleva asociados sus correspondientes multiplicadores.

Con esta terminología el teorema 9.13 se enuncia diciendo que los puntos regulares $\mathbf{p} \in M$ en los que la función diferenciable f presenta extremos relativos, condicionados por M, son necesariamente puntos estacionarios de $f|_M$.

En las aplicaciones más habituales todos los puntos de M son regulares y f es diferenciable en todos los puntos de M. En este caso cada extremo relativo de $f|_M$ se alcanza en un punto estacionario y la determinación de los extremos relativos de $f|_M$ comienza calculando los puntos estacionarios. Para ello se resuelve (cuando sea posible) el sistema de $n+m$ ecuaciones con $n+m$ incógnitas, x_1, x_2, ... x_n, μ_1, μ_2,.. μ_m

$$D_i f(x_1, x_2, \cdots x_n) - \textstyle\sum_{k=1}^m \mu_k D_i g_k(x_1, x_2, \cdots x_n) = 0, \ \ i = 1, 2 \cdots n$$

$$g_k(x_1, x_2, \cdots, x_n) = 0, \ \ k = 1, 2, \cdots m.$$

Cuando el sistema no tiene solución se puede asegurar que $f|_M$ no presenta extremos relativos. Si el sistema tiene un número finito de soluciones tendremos un conjunto finito de puntos estacionarios $E(f,M)$, acompañados sus correspondientes multiplicadores.

Si además M es compacto y el problema que nos ocupa consiste en determinar el máximo y el mínimo absoluto de f sobre M, bastará evaluar f en cada uno de los puntos estacionarios para obtener, por inspección directa, los extremos absolutos:

$$\max_{\mathbf{x} \in M} f(\mathbf{x}) = \max\{f(\mathbf{p}) : \mathbf{p} \in E(f,M)\}$$

$$\min_{\mathbf{x} \in M} f(\mathbf{x}) = \min\{f(\mathbf{p}) : \mathbf{p} \in E(f,M)\}$$

Lo mismo se puede hacer cuando M no sea compacto pero, por la naturaleza del problema, estemos seguros de que f alcanza en M un máximo, o un mínimo relativo, o ambos. Así por ejemplo la función

$$f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|^2 = (x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2$$

siempre alcanza un mínimo absoluto sobre M cuando $M \subset \mathbb{R}^n$ es cerrado y $\mathbf{a} \in \mathbb{R}^n$ (Recuérdese que la distancia de un punto $\mathbf{a} \in \mathbb{R}^n$ a un cerrado M, siempre se alcanza en algún $\mathbf{p} \in M$).

A la hora de las aplicaciones concretas, cuando f no es diferenciable en todos los puntos de M o existe algún punto de M donde $\mathbf{g}$ no es regular, para determinar los extremos absolutos de f se puede proceder como antes, pero añadiendo a los puntos estacionarios $E(f,M)$ el conjunto $S(f,M)$ formado por los puntos de M donde f no es diferenciable o $\mathbf{g}$ no es regular.

Un problema más delicado es el de determinar, entre los puntos estacionarios los que proporcionan un máximo relativo, o un mínimo relativo. En el siguiente teorema veremos condiciones de segundo orden suficientes para que un punto estacionario sea un punto de máximo relativo o un punto de mínimo relativo. Pero este criterio no es aplicable a los puntos de $S(f, M)$ ni a los puntos estacionarios donde f o $\mathbf{g}$ no sean diferenciables dos veces.

Dado un punto estacionario $\mathbf{p} \in E(f, M)$, para decidir si es punto de máximo o de mínimo relativo se puede aplicar el siguiente método: Se busca una parametrización de un entorno relativo $V_{\mathbf{p}} \cap M$, (donde $V_{\mathbf{p}} \subset \Omega$ es entorno abierto de $\mathbf{p}$), es decir un homeomorfismo de clase C^1, $\boldsymbol{\varphi} : U \to M \cap V_{\mathbf{p}}$, definido en un abierto $U \subset \mathbb{R}^k$, $k = m - n$, tal que $\mathbf{0} \in U$ y $\boldsymbol{\varphi}(\mathbf{0}) = \mathbf{p}$. Entonces $\mathbf{p}$ será punto de máximo (resp. mínimo) relativo para $f|_M$ si y sólo sí 0 es un punto de máximo (resp. mínimo) para la función real de k variables $F = f \circ \boldsymbol{\varphi}$, definida en U. A esta función F se le aplicarán con los métodos expuestos anteriormente para el caso de extremos ordinarios (sin ligaduras) y para ello será preciso calcular las derivadas segundas $D_{ij}F(\mathbf{p})$. El teorema de la función implícita asegura que existe esta parametrización local la cual, después de una permutación en las variables, se puede suponer de la forma $\boldsymbol{\varphi}(\mathbf{x}) = (\mathbf{x}, \mathbf{f}(\mathbf{x}))$. Es decir, en las ecuaciones de ligadura

$$g_i(x_1, x_2, \cdots x_n), \quad 1 \leq i \leq m$$

es posible despejar localmente m variables en función de las restantes. En general no es posible, o es bastante engorroso, obtener explícitamente las ecuaciones de las funciones implícitas. Sin embargo esto no supone una dificultad seria porque, aunque no conozcamos estas ecuaciones, es posible calcular, en el punto $\mathbf{p}$, las derivadas parciales primeras y segundas de las funciones implícitas. Usando la regla de la cadena, estas derivadas permiten calcular las derivadas segundas $D_{ij}F(\mathbf{p})$, que son las que intervienen en las condiciones de segundo orden expuestas en 6.17.

Teorema 9.11 *Sea $f : \Omega \to \mathbb{R}$ una función de clase $C^2(\Omega)$ definida en un abierto $\Omega \subset \mathbb{R}^n$ y $M \subset \Omega$ un conjunto de la forma*

$$M = \{\mathbf{x} \in \Omega : g_1(\mathbf{x}) = 0, g_2(\mathbf{x}) = 0, \cdots , g_m(\mathbf{x}) = 0\}$$

donde las funciones g_k, $1 \leq k \leq m$, son de clase $C^2(\Omega)$. Sea $\mathbf{p} \in M$ un punto estacionario de f sobre M, es decir, un punto donde los vectores $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots \nabla g_m(\mathbf{p})$ son linealmente independientes y existen coeficientes $\mu_1, \mu_2, \cdots \mu_m \in \mathbb{R}$ tales que

$$\nabla f(\mathbf{p}) = \mu_1 \nabla g_1(\mathbf{p}) + \mu_2 \nabla g_2(\mathbf{p}) + \cdots + \mu_m \nabla g_m(\mathbf{p})$$

Consideremos la función auxiliar $H_{\mathbf{p}}(\mathbf{x}) = f(\mathbf{x}) - \sum_{k=1}^{m} \mu_k g_k(\mathbf{x})$, y sea

$$Q_{\mathbf{p}}(\mathbf{u}) = d^2 H(\mathbf{p})\mathbf{u}^2 = \sum_{i,j=1}^{n} D_{ij} H_{\mathbf{p}}(\mathbf{p}) u_i u_j$$

la forma cuadrática asociada a esta función en el punto $\mathbf{p}$. Se verifica:

a) *Si* $f|_M$ *presenta en* $\mathbf{p}$ *un mínimo (resp. máximo) relativo entonces* $Q_{\mathbf{p}}(\mathbf{u}) \geq 0$ *(resp.* ≤ 0 *) para todo* $\mathbf{u} \in T_{\mathbf{p}}(M)$*.*

b) *Si* $Q_{\mathbf{p}}(\mathbf{u}) > 0$ *(resp.* < 0 *) para todo* $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$ *entonces* $f|_M$ *presenta en* $\mathbf{p}$ *un mínimo (resp. máximo) relativo estricto.*

DEM: Según el teorema de la función implícita la condición de que los vectores $\nabla g_1(\mathbf{p}), \nabla g_2(\mathbf{p}), \cdots, \nabla g_m(\mathbf{p})$ sean linealmente independientes garantiza la existencia de un entorno abierto de $\mathbf{p}$, $\Omega_{\mathbf{p}} \subset \Omega$, tal que $M \cap \Omega_{\mathbf{p}} = \boldsymbol{\varphi}(U)$, donde $\boldsymbol{\varphi} : U \to \mathbb{R}^n$ es una función de clase C^2 en un abierto $U \subset \mathbb{R}^k$, que establece un homeomorfismo entre U y $M \cap \Omega_{\mathbf{p}}$, con la propiedad de que los vectores

$$D_1\boldsymbol{\varphi}(\mathbf{s}), D_2\boldsymbol{\varphi}(\mathbf{s}), \cdots D_k\boldsymbol{\varphi}(\mathbf{s}) \in \mathbb{R}^n$$

son linealmente independientes para todo $\mathbf{s} \in U$. Podemos suponer que $\mathbf{0} \in U$ y que $\boldsymbol{\varphi}(\mathbf{0}) = \mathbf{p}$, con lo cual $T_{\mathbf{p}}(M) = d\boldsymbol{\varphi}(\mathbf{0})(\mathbb{R}^k)$. Como $\boldsymbol{\varphi} : U \to M \cap \Omega_{\mathbf{p}}$ es un homeomorfismo se sigue que $f|_M$ presenta en $\mathbf{p}$ un extremo relativo si y sólo si $F(\mathbf{s}) = f(\boldsymbol{\varphi}(\mathbf{s}))$ presenta en $\mathbf{s} = 0$ un extremo relativo del mismo tipo.

Cada vector $\mathbf{u} \in T_{\mathbf{p}}(M)$ es de la forma $\mathbf{u} = d\boldsymbol{\varphi}(\mathbf{0})\mathbf{h}$ con $\mathbf{h} \in \mathbb{R}^k$, y demostraremos que

$$d^2F(\mathbf{0})\mathbf{h}^2 = d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 \text{ para todo } \mathbf{u} = d\boldsymbol{\varphi}(\mathbf{0})\mathbf{h} \in T_{\mathbf{p}}(M)$$

Entonces aplicando los teoremas 6.11 y 6.15 a la función F, se obtendrá a) y b):

a) Si $f|_M$ presenta en $\mathbf{p}$ un mínimo relativo, entonces $d^2F(\mathbf{0})\mathbf{h}^2 \geq 0$ para todo $\mathbf{h} \in \mathbb{R}^k$, lo que equivale a que $d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 \geq 0$ para todo $\mathbf{u} \in T_{\mathbf{p}}(M)$.

b) Si $d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2 > 0$ para todo $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$, entonces $d^2F(\mathbf{0})\mathbf{h}^2 > 0$ para todo $\mathbf{h} \in \mathbb{R}^k \setminus \{\mathbf{0}\}$, luego F presenta en $\mathbf{s} = \mathbf{0}$ un mínimo relativo estricto y por lo tanto $f|_M$ también presenta en $\mathbf{p}$ un mínimo relativo estricto.

Si $(\varphi_1, \varphi_2, \cdots \varphi_n)$ son las componentes de $\boldsymbol{\varphi}$, las componentes de $\mathbf{u} = d\boldsymbol{\varphi}(\mathbf{0})\mathbf{h}$ vienen dadas por

$$u_k = d\varphi_k(\mathbf{0})\mathbf{h} = D_{\mathbf{h}}\varphi_k(\mathbf{0}), \quad k = 1, \cdots n$$

y la igualdad que queremos establecer, $d^2F(\mathbf{0})\mathbf{h}^2 = d^2H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2$, se escribe en la forma

$$D_{\mathbf{hh}}F(\mathbf{0}) = \sum_{k,j=1}^{n} D_{kj}H_{\mathbf{p}}(\mathbf{p})u_k u_j$$

Observemos en primer lugar que para todo $\mathbf{s} \in U$ y cada $1 \leq j \leq m$ se cumple $g_j(\boldsymbol{\varphi}(\mathbf{s})) = 0$, luego $F(\mathbf{s}) = f(\boldsymbol{\varphi}(\mathbf{s})) = H_{\mathbf{p}}(\boldsymbol{\varphi}(\mathbf{s}))$. Por consiguiente, según 7.3 tanto F como $H_{\mathbf{p}}$ son de clase C^2 en los abiertos donde están definidas.

Emprendemos el cálculo de $D_{\mathbf{hh}}F(\mathbf{0})$ a partir de la igualdad

$$dF(\mathbf{s}) = dH_{\mathbf{p}}(\boldsymbol{\varphi}(\mathbf{s})) \circ d\boldsymbol{\varphi}(\mathbf{s})$$

que se cumple para todo $\mathbf{s} \in U$. Si $\mathbf{h} \in \mathbb{R}^k$ resulta

$$D_{\mathbf{h}}F(\mathbf{s}) = dF(\mathbf{s})\mathbf{h} = dH_{\mathbf{p}}(\boldsymbol{\varphi}(\mathbf{s}))[d\boldsymbol{\varphi}(\mathbf{s})\mathbf{h}] = dH_{\mathbf{p}}(\boldsymbol{\varphi}(\mathbf{s}))D_{\mathbf{h}}\boldsymbol{\varphi}(\mathbf{s})$$

Existe $\delta > 0$ tal que si $t \in \mathbb{R}$ y $|t| < \delta$ entonces $t\mathbf{h} \in U$. Como $D_{\mathbf{h}}\varphi_i(\mathbf{s})$ son las componentes de $D_{\mathbf{h}}\boldsymbol{\varphi}(\mathbf{s})$, aplicando la igualdad anterior con $\mathbf{s} = t\mathbf{h}$ resulta:

$$D_{\mathbf{h}}F(t\mathbf{h}) = \sum_{j=1}^{n} D_j H_{\mathbf{p}}(\boldsymbol{\varphi}(t\mathbf{h}))D_{\mathbf{h}}\varphi_j(t\mathbf{h})$$

La derivada en $t = 0$ de la función en el lado izquierdo de la igualdad anterior es $D_{\mathbf{hh}}F(\mathbf{0})$. Por otra parte, la derivada de la función en el lado derecho de la igualdad se puede calcular usando la regla para derivar un producto y teniendo presente que la derivada en $t = 0$ de $D_j H_{\mathbf{p}}(\boldsymbol{\varphi}(t\mathbf{h}))$ viene dada por $\sum_{k=1}^{n} D_k D_j H_{\mathbf{p}}(\boldsymbol{\varphi}(0))D_{\mathbf{h}}\varphi_k(0)$. Se obtiene así

$$D_{\mathbf{hh}}F(\mathbf{0}) = \sum_{j=1}^{n}\left[\sum_{k=1}^{n} D_{kj}H_{\mathbf{p}}(\mathbf{p})D_{\mathbf{h}}\varphi_k(\mathbf{0})\right]D_{\mathbf{h}}\varphi_j(\mathbf{0}) + \sum_{j=1}^{n} D_j H_{\mathbf{p}}(\mathbf{p})D_{\mathbf{hh}}\varphi_j(\mathbf{0})$$

y sustituyendo $D_{\mathbf{h}}\varphi_k(\mathbf{0}) = u_k$, resulta

$$D_{\mathbf{hh}}F(\mathbf{0}) = \sum_{k,j=1}^{n} D_{kj}H_{\mathbf{p}}(\mathbf{p})u_k u_j + \sum_{j=1}^{n} D_j H_{\mathbf{p}}(\mathbf{p})D_{\mathbf{hh}}\varphi_j(\mathbf{0})$$

En virtud de su definición la función $H_{\mathbf{p}}$ cumple $D_j H_{\mathbf{p}}(\mathbf{p}) = 0$, si $1 \leq j \leq n$, luego

$$D_{\mathbf{hh}}F(\mathbf{0}) = \sum_{k,j=1}^{n} D_{k,j}H_{\mathbf{p}}(\mathbf{p})u_k u_j = d^2 H_{\mathbf{p}}(\mathbf{p})\mathbf{u}^2$$

■

9.3. Ejercicios resueltos

Ejercicio 9.12 *Compruebe que la curva*

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

es una subvariedad de $\mathbb{R}^3$, de clase C^∞ y dimensión 1. Obtenga los puntos de C donde la recta tangente es horizontal (paralela al plano $z = 0$).

SOLUCIÓN

Como las funciones $g_1(x, y, z) = x^2 + y^2 + z^2 - 1$, $g_2(x, y, z) = x^2 + 2y^2 - 2z$, son de clase C^∞, basta ver que para todo $(x, y, z) \in C$ son linealmente independientes los vectores $\nabla g_1(x, y, z) = (2x, 2y, 2z)$, $\nabla g_2(x, y, z) = (2x, 4y, -2)$, o lo que es lo mismo, que la matriz

$$\begin{pmatrix} x & y & z \\ x & 2y & -1 \end{pmatrix}$$

tiene rango 2 para cada punto $(x, y, z) \in C$.

El determinante formado con las dos primeras columnas no es nulo si $xy \neq 0$. Cuando $x = 0$, en los puntos de la forma $(0, y, z) \in C$ debe ser $y \neq 0$ y entonces no es nulo el determinante formado con las dos últimas columnas, que vale $-y(1+2z)$ (obsérvese que $z \geq 0$ para todo $(x, y, z) \in C$). Análogamente, cuando $y = 0$, en los puntos $(x, 0, z) \in C$ debe ser $x \neq 0$, y entonces no es nulo el determinante formado con las columnas primera y tercera.

Obtengamos los puntos de C donde la recta tangente es horizontal:
En un punto genérico $\mathbf{p} = (x_0, y_0, z_0) \in C$, los vectores $\nabla g_1(\mathbf{p})$, $\nabla g_2(\mathbf{p})$ son ortogonales, respectivamente, a los planos tangentes en ese punto a las superficies que determinan la curva, $g_1(x, y, z) = 0$, $g_2(x, y, z) = 0$. La recta tangente a C en $\mathbf{p}$ es la intersección de estos dos planos, y así, con el producto vectorial $\nabla g_1(\mathbf{p}) \times \nabla g_2(\mathbf{p})$ se consigue un vector de dirección de la recta:

$$(-y_0(1 + 2z_0), x_0(1 + z_0), x_0y_0)$$

La recta tangente a C en $\mathbf{p}$ es horizontal cuando la tercera componente de este vector es nula, es decir, cuando $x_0y_0 = 0$, y esto ocurre en los puntos donde la curva corta a los planos $x = 0$, e $y = 0$. Estos puntos se calculan resolviendo primero el sistema de dos ecuaciones con incógnitas y, z, que resulta al sustituir $x = 0$ en las ecuaciones de las superficies, y luego el sistema de dos ecuaciones con las incógnitas x, z obtenido al sustituir $y = 0$ en dichas ecuaciones. ■

Ejercicio 9.13 *Calcule los puntos de la hipérbola* $xy = 1$ *que están más cerca del punto* $(2, 1)$ *(con la distancia usual).*

SOLUCIÓN

La distancia del punto $(2, 1)$ a la hipérbola $M = \{(x, y) \in \mathbb{R}^2 : xy - 1 = 0\}$ se alcanza en algún punto $(a, b) \in M$, ya que M es un subconjunto cerrado de $\mathbb{R}^2$.

El punto (a, b) que minimiza la distancia de $(2, 1)$ a la hipérbola también minimiza su cuadrado $f(x, y) = (x - 2)^2 + (y - 1)^2$, que es una función más cómoda de manejar. Observemos en primer lugar que en cada $(x, y) \in M$ la condición de ligadura $g(x, y) = xy - 1$ tiene gradiente no nulo $\nabla g(x, y) = (y, x)$. luego el punto (a, b) que andamos buscando se encontrará entre los puntos estacionarios de f sobre M. Estos puntos son soluciones del sistema de tres ecuaciones con tres incógnitas

$$D_1f(x, y) - \mu D_1g(x, y) = 0, \quad D_2f(x, y) - \mu D_2g(x, y) = 0, \quad xy - 1 = 0$$

que se concretan en

$$x - \mu y = 2, \qquad -\mu x + y = 1, \quad xy = 1$$

Multiplicando las dos primeras ecuaciones por x, y usando la tercera ecuación se elimina la variable y y se obtienen las dos ecuaciones

$$\mu = x^2 - 2x, \quad \mu x^2 = 1 - x$$

Eliminando μ, se llega a la ecuación $x^4 - 2x^3 + x - 1 = 0$ con dos soluciones reales, que se pueden calcular con un programa de cálculo simbólico (p.e DERIVE)

$$x_1 = \frac{1}{2} - \sqrt{\sqrt{5}/2 + 3/4}; \quad x_2 = \frac{1}{2} + \sqrt{\sqrt{5}/2 + 3/4}$$

Sus valores aproximados, $x_1 = -0{,}866760$, $x_2 = 1{,}86676$, también se pueden calcular con los métodos habituales de cálculo numérico. Son las abscisas de dos puntos, uno en cada rama de la hipérbola, donde la distancia alcanza mínimos relativos. Es claro que el mínimo absoluto se alcanza en el punto $(a, b) = (x_2, 1/x_2)$, situado en la rama del primer cuadrante. ■

Ejercicio 9.14 *Utilice el método de los multiplicadores de Lagrange para calcular las dimensiones de la caja de superficie mínima que encierra un volumen de 1 litro.*

solución

Si $x > 0$, $y > 0$, $z > 0$ son las dimensiones de la caja, expresadas en cm., se trata de minimizar el área $S(x, y, z) = 2xy + 2xz + 2yz$ cuando las variables x, y, z están sometidas a la condición de que el volumen encerrado sea $xyz = 1000$ cm^3.

Este problema ya fue resuelto en 5.39 convirtiéndolo en un problema de extremos ordinarios para la función de dos variables $f(x, y) = S(x, y, 1000/(xy))$, sobre el abierto $U = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0\}$. Ahora se trata de resolverlo como un problema de extremos condicionados, para la función $S(x, y, z) = 2xy + 2xz + 2yz$ sobre la superficie

$$M = \{(x, y, z) \in \Omega : xyz = 1000\}, \quad \text{donde} \quad \Omega = \{(x, y, z) : x > 0, y > 0, z > 0\}$$

Un razonamiento similar al efectuado en 5.39 permite justificar que $S|_M$ alcanza un mínimo absoluto: El trozo de superficie

$$K = \{(x, y, z) \in M : x \geq 1, \ y \geq 1, \ z \geq 1\}$$

es cerrado y acotado (pues $M \subset [1, 1000]^3$) y por lo tanto compacto.

Cuando $(x, y, z) \in M \setminus K$ alguna de sus componentes es menor que 1, y si suponemos que $x < 1$ se tendrá $2yz = 2000/x > 2000$, luego $S(x, y, z) > 2000$.

Como existe $\mathbf{p} \in K$ con $S(\mathbf{p}) < 2000$, podemos asegurar que el mínimo absoluto de S sobre el compacto K también es el mínimo absoluto de $S|_M$.

Para cada $(x, y, z) \in M$, la función $g(x, y, z) = xyz - 1000$ cumple

$$\nabla g(x, y, z) = (yz, xz, xy) \neq (0, 0, 0)$$

luego, en virtud de 9.10, el mínimo absoluto de $S|_M$ se alcanza en uno de los puntos estacionarios de $S|_M$, es decir, en una de las soluciones del sistema de ecuaciones

$$2y + 2z - \mu yz = 0; \quad 2x + 2z - \mu xz = 0; \quad 2y + 2x - \mu xy = 0; \quad xyz = 1000.$$

Multiplicando la primera ecuación por $x > 0$, la segunda por $y > 0$ y la tercera por $z > 0$ resulta

$$2xy + 2xz = 2000\mu; \quad 2xy + 2zy = 2000\mu; \quad 2yz + 2xz = 2000\mu;$$

cuya única solución en Ω es $x = y = z = 10$, $\mu = 2/5$. La unicidad de la solución permite afirmar que el mínimo absoluto de $S|_M$ se alcanza cuando $x = y = z = 10$.
NOTA: Aunque ya sabemos que en $\mathbf{p} = (10, 10, 10)$ hay un mínimo absoluto, a título ilustrativo comprobaremos que se cumple la condición suficiente de mínimo relativo vista en el teorema 9.11: Para ello formamos la función auxiliar

$$H(x, y, z) = S(x, y, z) - \frac{2}{5}(xyz - 1000)$$

Su matriz hessiana en $(10, 10, 10)$ es

$$\begin{pmatrix} 0 & -2 & -2 \\ -2 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix}$$

y la forma cuadrática asociada a esta matriz es

$$Q_{\mathbf{p}}(\mathbf{u}) = \sum_{i,j=1}^{n} D_{ij}H(\mathbf{p})u_i u_j = -4(u_1u_2 + u_1u_3 + u_2u_3)$$

Tenemos que comprobar que $Q_{\mathbf{p}}(\mathbf{u}) > 0$ para todo $\mathbf{u} \in T_{\mathbf{p}}(M) \setminus \{\mathbf{0}\}$.
Efectivamente, la ecuación implícita del plano tangente $T_{\mathbf{p}}(M)$ es $u_1 + u_2 + u_3 = 0$, y para restringir la forma cuadrática a los vectores de este plano sustituimos $u_3 = -u_1 - u_2$ y así obtenemos una forma cuadrática en dos variables

$$q(u_1, u_2) = -4(u_1u_2 - (u_1 + u_2)^2) = 4(u_1^2 + u_2^2 + u_1u_2) = (u_1 + \frac{1}{2}u_2)^2 + \frac{3}{4}u_2^2$$

luego $Q(u_1, u_2, yu_3) = q(u_1, u_2) > 0$ cuando $u_1+u_2+u_3 = 0$, y $(u_1, u_2, u_3) \neq (0, 0, 0)$.
■

Ejercicio 9.15 *Compruebe que* $\mathbf{p} = (1, 1, 2)$ *es un punto estacionario de la función* $f(x, y, z) = (2 - x)yz$ *sobre la superficie* $M = \{(x, y, z) : 8x - 4y^2 - z^2 = 0\}$, *y que en este punto* $f|_M$ *presenta un máximo relativo.*

SOLUCIÓN

Si $g(x, y, z) = 8x - 4y^2 - z^2$, se tiene $\nabla g(\mathbf{p}) = (8, -8, -4)$, y $\nabla f(\mathbf{p}) = (-2, 2, 1)$, luego $\nabla f(\mathbf{p}) = -\frac{1}{4}\nabla g(\mathbf{p})$. Por lo tanto $\mathbf{p}$ es un punto estacionario de $f|_M$ con multiplicador $\mu = -1/4$. Para estudiar la naturaleza de este punto estacionario consideramos la función $H(x, y, z) = f(x, y, z) - \frac{1}{4}g(x, y, z)$, cuya matriz Hessiana en el punto $\mathbf{p}$ es

$$\begin{pmatrix} 0 & -2 & -1 \\ -2 & -2 & 1 \\ -1 & 1 & -1/2 \end{pmatrix}$$

La forma cuadrática asociada a esta matriz es

$$Q(u_1, u_2, u_3) = -2u_2^2 - \frac{1}{2}u_3^2 - 4u_1u_2 - 2u_1u_3 + 2u_2u_3$$

Esta forma cuadrática la tenemos que restringir al plano tangente $T_{\mathbf{p}}(M)$ cuya ecuación es $2u_1 - 2u_2 - u_3 = 0$. Sustituyendo $u_3 = 2(u_1 - u_2)$ obtenemos

$$q(u_1, u_2) = Q(u_1, u_2, 2(u_1 - u_2)) = -6u_1^2 - 8u_2^2 + 8u_1u_2 = Au_1^2 + 2Bu_1u_2 + Cu_2$$

donde $A = -6$, $B = 4$ y $C = -8$. Como $AC - B^2 > 0$ y $A < 0$, la forma cuadrática $q(u_1, u_2)$ es definida negativa, luego $Q(\mathbf{u}) = q(u_1, u_2) < 0$ para cada vector tangente no nulo $\mathbf{u} \in T_{\mathbf{p}}(M)$. Entonces, según el teorema 9.11, podemos afirmar que $f|_M$ presenta un máximo relativo en $\mathbf{p} = (1, 1, 2)$.
Segunda solución. También se puede justificar la última afirmación usando la técnica de la función implícita: Si $z = z(x, y)$ es la función implícita definida por $8x - 4y^2 - z^2 = 0$ en un entorno de $(1, 1, 2)$, se considera la función compuesta $F(x, y) = f(x, y, z(x, y)$ y se comprueba que F presenta un máximo relativo en el punto $(x_0, y_0) = (1, 1)$. Para calcular las derivadas parciales primeras y segundas de F en $(1, 1)$, debemos comenzar calculando

$$z_x(1, 1), \quad z_y(1, 1), \quad z_{xy}(1, 1) = z_{yx}(1, 1), \quad z_{yy}(1, 1)$$

Aunque en el caso que nos ocupa tenemos una fórmula concreta para la función implícita, $z = 2\sqrt{2x - y^2}$, preferimos efectuar el cálculo de estas derivadas parciales sin usar la fórmula, mediante la técnica usual de derivación de funciones implícitas: Se calculan las derivadas parciales sucesivas, respecto a x y respecto a y, en la identidad $8x - 4y^2 - z(x, y)^2 = 0$, y se sustituyen los valores concretos $x = 1$, $y = 1$, $z = 2$. (Por comodidad escribimos z, z_x, z_y, z_{xx}..., omitiendo el punto (x, y) en el que se evalúan estas funciones). En lo que sigue la notación $\rightarrow$ indica que el resultado escrito a la derecha de $\rightarrow$ se ha obtenido con la sustitución mencionada en el término de la izquierda.

$$\begin{array}{rrcl}
\text{i)} & 8 - 2zz_x = 0; & \rightarrow & z_x(1, 1) = 2. \\
\text{ii)} & -8y - 2zz_y = 0; & \rightarrow & z_y(1, 1) = -2.
\end{array}$$

Derivando respecto a x y respecto a y en i) y derivando respecto a y en ii)se obtiene

$$\begin{array}{rrcl}
\text{iii)} & (z_x)^2 + zz_{xx} = 0; & \rightarrow & z_{xx}(1, 1) = -2. \\
\text{iv)} & z_yz_x + zz_{xy} = 0; & \rightarrow & z_{xy}(1, 1) = z_{yx}(1, 1) = -2. \\
\text{v)} & -8 - 2(z_y)^2 - 2zz_{yy} = 0; & \rightarrow & z_{yy}(1, 1) = -4.
\end{array}$$

Utilizando la regla de la cadena

$$D_1F(x, y) = -yz + (2 - x)yz_x; \quad D_2F(x, y) = (2 - x)z + (2 - x)yz_y$$

y sustituyendo $x = 1$, $y = 1$, $z = 2$, se llega a los valores

$$D_1F(1, 1) = -2 + 2 = 0; \quad D_2F(1, 1) = 2 - 2 = 0,$$

luego $(1,1)$ es un punto estacionario de F. Para ver que F presenta un máximo relativo en este punto necesitamos calcular

$$A = D_{11}F(1,1); \quad B = D_{12}F(1,1) = D_{21}F(1,1); \quad C = D_{22}F(1,1).$$

$$\begin{array}{lll}
D_{11}F = -2yz_x + (2-x)yz_{xx}; & \to & A = D_{11}F(1,1) = -6; \\
D_{21}F = -z - yz_y + (2-x)(z_x + yz_{xy}); & \to & B = D_{21}F(1,1) = 4; \\
D_{22}F = (2-x)(2z_y + yz_{yy}); & \to & C = D_{22}F(1,1) = -8;
\end{array}$$

Como $AC - B^2 > 0$ y $A < 0$ se sigue que F presenta un máximo relativo en $(1,1)$.

Tercera solución. Se puede ver directamente que $f|_M$ presenta en $\mathbf{p}$ un máximo relativo: Lo haremos viendo que hay un entorno relativo $M_\mathbf{p}$ de $\mathbf{p}$ en M que cumple
a) $f|_{M_\mathbf{p}}$ alcanza un máximo absoluto en algún $\mathbf{q} \in M_\mathbf{p}$
b) $\mathbf{p}$ es el único punto estacionario de $f|_{M_\mathbf{p}}$
ya que entonces, al ser $\mathbf{p}$ y $\mathbf{q}$ puntos estacionarios de $f|_{M_\mathbf{p}}$, deberá ser $\mathbf{p} = \mathbf{q}$.

Veamos que $M_\mathbf{p} = \{(x,y,z) \in M : \ 0 < x < 2, \ 0 < y, \ 0 < z\}$ cumple estas condiciones.
a) El conjunto $K = \{(x,y,z) : 8x - 4y^2 - z^2 = 0, 0 \leq x \leq 2, \ 0 \leq y, \ 0 \leq z\}$ es compacto (por ser cerrado y acotado), luego existe $\mathbf{q} \in K$ tal que

$$f(\mathbf{q}) = \max\{f(x,y,z) : (x,y,z) \in K\}$$

Es claro que $f(\mathbf{q}) > 0$ (porque hay puntos en K, como $(5/8,1,1) \in K$, con $f(5/8,1,1) > 0$) y también es evidente que $f(x,y,z) = 0$ cuando $(x,y,z) \in K$ cumple alguna de las igualdades $y = 0$, $z = 0$, $x = 2$, $x = 0$ (obsérvese que $x = 0 \Rightarrow 4y^2 + z^2 = 0$, luego $y = 0$ y $z = 0$). Se sigue de esto que $\mathbf{q} \in M_\mathbf{p} \subset K$, luego $f(\mathbf{q})$ es el máximo absoluto de $f|_{M_\mathbf{p}}$.

Según el método de los multiplicadores de Lagrange los puntos estacionarios de f sobre $M_\mathbf{p}$ se obtienen calculando las soluciones (x,y,z,μ) del sistema

$$\begin{array}{l}
-yz - 8\mu = 0; \\
(2-x)z + 8\mu y = 0; \\
(2-x)y + 2\mu z = 0; \\
8x - 4y^2 - z^2 = 0;
\end{array}$$

que pertenecen a $M_\mathbf{p}$ y es fácil ver que $x = 1$, $y = 1$, $z = 2$, $\mu = -1/4$ es la única solución del sistema que cumple esta condición.

Efectivamente, multiplicando la segunda ecuación por y y la tercera por z se obtiene $8\mu y^2 = 2\mu z^2$. Como buscamos soluciones con $y > 0$, $z > 0$ podemos asegurar, en virtud de la primera ecuación, que $\mu \neq 0$, luego $8y^2 = 2z^2$ y por lo tanto $2y = z$. Sustituyendo $z = 2y$ en la primera y en la última ecuación obtenemos $8\mu = -2y^2$, $x = y^2$. Llevando estos valores a la tercera ecuación se llega a la ecuación $0 = (2 - y^2)2y - 2y^3 = 4(y - y^3)$, cuya única solución con $y > 0$ es $y = 1$, luego $\mathbf{p} = (1,1,2)$ es el único punto estacionario de $F|_{M_\mathbf{p}}$. ■

Ejercicio 9.16 *Determine los valores de los parámetros a, b para los que la función*

$$f(x,y,z) = x^2 + y^2 + 2axy + 2bz$$

presenta en $\mathbf{p} = (1,1,1)$ un máximo relativo sobre la esfera $x^2 + y^2 + y^2 = 3$.

SOLUCIÓN

La función $g(x,y,z) = x^2 + y^2 + z^2 - 3 = 0$ cumple que el vector $\nabla g(\mathbf{p}) = (2,2,2)$ no es nulo. Si f presenta un extremo relativo sobre la esfera en $\mathbf{p}$, según la condición necesaria de extremo condicionado, debe existir $\mu \in \mathbb{R}$ verificando:
$0 = D_1 f(1,1,1) - \mu D_1 g(1,1,1) = 2 + 2a - 2\mu$,
$0 = D_2 f(1,1,1) - \mu D_2 g(1,1,1) = 2 + 2a - 2\mu b$,
$0 = D_2 f(1,1,1) - \mu D_2 g(1,1,1) = 2b - 2\mu b$
de donde se obtiene que $\mu = b = a + 1$. Por lo tanto, en lo que sigue suponemos que $b = a + 1$, y así tenemos garantizado que $\mathbf{p}$ es un punto estacionario de f sobre la esfera. Para discutir cuando este punto estacionario es un punto de máximo relativo condicionado, consideramos la función

$$H(x,y,z) = f(x,y,z) - b(x^2 + y^2 + z^2 - 3)$$

cuya matriz Hessiana en el punto $\mathbf{p}$ es

$$\begin{pmatrix} 2(1-b) & 2(b-1) & 0 \\ 2(b-1) & 2(1-b) & 0 \\ 0 & 0 & -2b \end{pmatrix}$$

luego, la forma cuadrática asociada es

$$Q(u_1, u_2.u_3) = 2(b-1)(u_1^2 + u_2^2 - 2u_1u_2) - 2bu_3^2$$

Una condición suficiente para que f presente en $\mathbf{p}$ un máximo relativo condicionado es que esta forma cuadrática sea definida negativa sobre el plano tangente a la esfera en $\mathbf{p}$, cuya ecuación es $u_1 + u_2 + u_3 = 0$. Sustituyendo $u_3 = -(u_1 + u_2)$ se obtiene una forma cuadrática en dos variables

$$q(u_1, u_2) = Q(u_1, u_2, -(u_1 + u_2)) = Au_1^2 + 2Bu_1u_2 + Cu_2^2$$

donde $A = C = 2 - 4b$, $B = 2$, luego $AC - B^2 = 16b(b-1)$.

Si $b > 1$ es $AC - B^2 > 0$ y $A < 0$, luego la forma cuadrática $q(u_1, u_2)$ es definida negativa, lo que significa que $Q(\mathbf{u}) < 0$ para cada cada vector no nulo $\mathbf{u}$ tangente a la esfera en $\mathbf{p}$, luego $\mathbf{p}$ es un punto de máximo relativo condicionado.

Cuando $0 < b < 1$ se cumple $AC - B^2 < 0$ y por lo tanto la forma cuadrática $q(u_1, u_2)$ es indefinida, lo que significa que existen $\mathbf{u}, \mathbf{v}$, vectores tangentes en $\mathbf{p}$ a la esfera tales que $Q(\mathbf{u}) < 0 < Q(\mathbf{v})$, luego, en este caso en el punto $\mathbf{p}$ no hay extremo condicionado.

Cuando $b < 0$ se cumple $AC - B^2 > 0$, y $A > 0$ luego la forma cuadrática $q(u_1, u_2)$ es definida positiva, luego $Q(\mathbf{u}) > 0$ para cada cada vector no nulo $\mathbf{u}$ tangente a la esfera en $\mathbf{p}$, y por lo tanto $\mathbf{p}$ es un punto de mínimo relativo condicionado.

Falta discutir lo que ocurre cuando $b = 0$ y cuando $b = 1$.

Si $b = 0$ es $a = -1$ y es claro que $f(x, y, z) = (x - y)^2$ no presenta en $\mathbf{p}$ un máximo relativo condicionado.

Si $b = 1$ es $a = 0$, y ahora la función $f(x, y, z) = x^2 + y^2 + 2z$, al restringirla a la esfera sólo depende de z: $f(x, y, z) = 3 - z^2 + 2z$. Como $3 - z^2 + 2z$ presenta en $z = 1$ un máximo relativo, se sigue que, en este caso, $f(x, y, z) = x^2 + y^2 + 2z$ presenta en el punto $(1, 1, 1)$. un máximo relativo condicionado.

Segunda solución: La discusión anterior también se puede realizar con la técnica de la función implícita, estudiando cuando $(1, 1)$ es un punto de máximo relativo ordinario para la función de dos variables reales

$$F(x, y) = f(x, y, z(x, y)) = x^2 + y^2 + 2axy + 2bz(x, y)$$

donde $z(x, y)$ es la función implícita que define la ecuación $g(x, y, z) = 0$ en un entorno del punto $(1, 1, 1)$. (Aunque en este caso tenemos una fórmula explícita $z(x, y) = \sqrt{3 - x^2 - y^2}$, haremos las cuentas que siguen, sin usarla).

Para este estudio necesitamos calcular la matriz Hessiana de F en $(1, 1)$.
$F_x(x, y) = 2x + 2ay + 2bz_x(x, y)$; $F_y(x, y) = 2y + 2ax + 2bz_x(x, y)$;
$F_{xx}(x, y) = 2 + 2bz_{xx}(x, y)$; $F_{yy}(x, y) = 2 + 2bz_{yy}(x, y)$;
$F_{xy}(x, y) = 2a + 2bz_{xy}(x, y)$; $F_{yx}(x, y) = 2a + 2bz_{xy}(x, y)$;
Calculamos las derivadas parciales $z_{xx}(1, 1)$, $z_{yy}(1, 1)$, $z_{xy}(1, 1)$, con la técnica de derivación implícita: Derivando respecto a x en la identidad $x^2 + y^2 + z(x, y)^2 = 3$ resulta $2x + 2z(x, y)z_x(x, y) = 0$ y cuando $x = y = z = 1$ se obtiene $z_x(1, 1) = -1$. Análogamente se calcula $z_y(1, 1) = -1$. Derivando respecto a x y respecto a y en la identidad $2x + 2z(x, y)z_x(x, y) = 0$, y sustituyendo los valores particulares $x = y = z = 1$ se obtiene $z_{xx}(1, 1) = -2$. Análogamente $z_{xy}(1, 1) = z_{yx}(1, 1) = -1$, y $z_{yy}(1, 1) = -2$. Sustituyendo arriba estos valores resulta:

$$F_{xx}(1, 1) = F_{yy}(1, 1) = 2 - 4b, \quad F_{xy}(1, 1) = F_{yx}(1, 1) = -2$$

luego el determinante Hessiano de F en $(1, 1)$ vale

$$\Delta(b) = \begin{vmatrix} 2 - 4b & -2 \\ -2 & 2 - 4b \end{vmatrix} = 16b(b - 1)$$

y se acaba la discusión como antes. ■

Ejercicio 9.17 *Sea $Q(x, y) = Ax^2 + 2Bxy + Cy^2$ una forma cuadrática no idénticamente nula. Demuestre que*

$$m_1 = \text{máx}\{Q(x, y) : x^2 + y^2 = 1\} \quad y \quad m_2 = \text{mín}\{Q(x, y) : x^2 + y^2 = 1\}$$

son las soluciones de la ecuación de segundo grado $\begin{vmatrix} A - \mu & B \\ B & C - \mu \end{vmatrix} = 0.$

Deduzca de ello:

i) Q es indefinida si y sólo si $AC - B^2 < 0$.

ii) Q es definida positiva si y sólo si $AC - B^2 > 0$ y $A > 0$.

iii) Q es definida negativa si y sólo si $AC - B^2 > 0$ y $A < 0$.

Si $\mathbf{u}_1 = (x_1, y_1)$, y $\mathbf{u}_2 = (x_2, y_2)$ son puntos de la circunferencia $x^2 + y^2 = 1$ donde Q alcanza el máximo y el mínimo absoluto, $m_1 = Q(x_1, y_1)$, $m_2 = Q(x_2, y_2)$, demuestre que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales. ¿Cuál es la interpretación geométrica de estos resultados.?

(Indicación : La aplicación lineal $L(x, y) = (Ax + By, Bx + Cy)$ es simétrica, $\langle\ L(\mathbf{w}) \mid \mathbf{w}\ \rangle = \langle \mathbf{w} \mid L(\mathbf{w})\ \rangle$, y verifica $Q(\mathbf{w}) = \ \langle\ \mathbf{w} \mid L(\mathbf{w})\ \rangle$.)

solución

$M = \{(x, y) : x^2 + y^2 - 1 = 0\}$ es compacto, luego existen $(x_1, y_1) \in M$, $(x_2, y_2) \in M$ donde la función continua Q alcanza sus extremos absolutos:

$$m_1 = Q(x_1, y_1), \quad m_2 = Q(x_2, y_2).$$

En todo $(x, y) \in M$ la función $g(x, y) = x^2 + y^2 - 1$ cumple $\nabla g(x, y) \neq (0, 0)$, luego (x_1, y_1), (x_2, y_2) son puntos estacionarios de $Q|_M$, es decir, existen $\mu_1, \mu_2 \in \mathbb{R}$ tales que (x_1, y_1, μ_1) , (x_2, y_2, μ_2) son soluciones del sistema de tres ecuaciones

$$D_1Q(x, y) - \mu D_1 g(x, y) = 0; \quad D_2Q(x, y) - \mu D_2 g(x, y) = 0; \quad g(x, y) = 0;$$

que se concretan en

$$Ax + By = \mu x; \quad Bx + Cy = \mu y; \quad x^2 + y^2 = 1.$$

luego toda solución (x, y, μ) de este sistema cumple $L(x, y) = \mu(x, y)$.

Multiplicando la primera ecuación por x, la segunda por y, y utilizando la tercera ecuación se obtiene que también se cumple $Q(x, y) = \mu$. En particular

a) $L(x_1, y_1) = \mu_1(x_1, y_1)$; $L(x_2, y_2) = \mu_2(x_2, y_2)$;

b) $m_1 = Q(x_1, y_1) = \mu_1$; $m_2 = Q(x_2, y_2) = \mu_2$.

Cuando $\mu = \mu_1$ y cuando $\mu = \mu_2$, el sistema lineal

$$(A - \mu)x + By = 0; \quad Bx + (C - \mu)y = 0;$$

tiene, respectivamente, las soluciones no triviales (x_1, y_1), (x_2, y_2), luego su determinante es nulo, luego μ_1, μ_2 son las soluciones de la ecuación del enunciado:

$$\mu^2 - (A + C)\mu + (AC - B^2) = 0.$$

De las dos soluciones reales de esta ecuación la mayor es el máximo $\mu_1 = m_1$, y la menor el mínimo $\mu_2 = m_2$. Es claro que

i) Q es indefinida si y sólo si $m_2 < 0 < m_1$.

ii) Q es definida positiva si y sólo si $m_2 > 0$.

iii) Q es definida negativa si y sólo si $m_1 < 0$.

Teniendo en cuenta que $m_1 + m_2 = A + C$, $m_1 m_2 = AC - B^2$, se obtiene:
a) $AC - B^2 < 0$ si y sólo si m_1 y m_2 tienen distinto signo, lo que ocurre si y sólo si Q es indefinida.
b) $AC - B^2 > 0$ si y sólo si m_1 y m_2 tienen el mismo signo, lo que ocurre si y sólo si Q es definida positiva o definida negativa. Q será definida positiva (resp. negativa) cuando $m_1 + m_2 = A + C > 0$ (resp. < 0) lo que ocurre si y sólo sí $A > 0$ (resp. $A < 0$). Obsérvese que A y C tienen el mismo signo porque $AC > B^2 \geq 0$.

Para lo que sigue, consideramos una base ortonormal de $\mathbb{R}^2$, $(\mathbf{v}_1, \mathbf{v}_2)$ con $\mathbf{v}_1 = \mathbf{u}_1$. Para demostrar que $\langle \mathbf{u}_1 \mid \mathbf{u}_2 \rangle = 0$ basta ver que cuando se expresa $\mathbf{u}_2$ respecto a esta base, $\mathbf{u}_2 = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2$, se cumple que $\alpha = 0$. Obsérvese que $\alpha^2 + \beta^2 = \langle \mathbf{u}_2 \mid \mathbf{u}_2 \rangle = 1$.
Utilizando que $Q(\mathbf{w}) = \langle \mathbf{w} \mid L(\mathbf{w}) \rangle$ y la bilinealidad del producto escalar

$$Q(\mathbf{u}_2) = \langle \mathbf{u}_2 \mid L(\mathbf{u}_2) \rangle = \langle \alpha \mathbf{v}_1 + \beta \mathbf{v}_2 \mid \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2) \rangle =$$

$$= \alpha^2 \langle \mathbf{v}_1 \mid L(\mathbf{v}_1) \rangle + \alpha\beta \langle \mathbf{v}_1 \mid L(\mathbf{v}_2) \rangle + \alpha\beta \langle \mathbf{v}_2 \mid L(\mathbf{v}_1) \rangle + \beta^2 \langle \mathbf{v}_2 \mid L(\mathbf{v}_2) \rangle$$

Utilizando la simetría de L y que $L(\mathbf{v}_1) = \mu_1 \mathbf{v}_1$, (véase a)) resulta

$$\langle \mathbf{v}_1 \mid L(\mathbf{v}_2) \rangle = \langle L(\mathbf{v}_1) \mid \mathbf{v}_2 \rangle = \langle \mathbf{v}_2 \mid L(\mathbf{v}_1) \rangle = \mu_1 \langle \mathbf{v}_2 \mid \mathbf{v}_1 \rangle = 0$$

luego

$$Q(\mathbf{u}_2) = \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_2)$$

De aquí se deduce que $Q(\mathbf{v}_2) < Q(\mathbf{v}_1)$: Sabemos que $Q(\mathbf{v}_2) \leq m_1 = Q(\mathbf{v}_1)$, pero no se puede dar la igualdad porque en ese caso se tendría

$$m_2 = Q(\mathbf{u}_2) = \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_1) = Q(\mathbf{v}_1) = m_1$$

y Q sería constante.
Con la desigualdad $Q(\mathbf{v}_2) < Q(\mathbf{v}_1)$ se obtiene que $\alpha = 0$: Si fuese $\alpha \neq 0$ se tendría $\alpha^2 Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1)$ es decir $(1 - \beta^2) Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1)$, luego

$$Q(\mathbf{v}_2) < \alpha^2 Q(\mathbf{v}_1) + \beta^2 Q(\mathbf{v}_2) = Q(\mathbf{u}_2)$$

y esta desigualdad es falsa porque $Q(\mathbf{u}_2)$ es el mínimo absoluto de Q sobre M.

Interpretación geométrica:
Para cada $c > 0$ las curvas de nivel $N_c = \{(x, y) : Q(x, y) = c\}$ son simétricas respecto al origen. Obtengamos la ecuación de N_c respecto a la base ortonormal $\{\mathbf{u}_1, \mathbf{u}_2\}$. Si (s, t) son las coordenadas de (x, y) respecto a esta base, usando que $L(\mathbf{u}_1) = \mu_1 \mathbf{u}_1$ y $L(\mathbf{u}_2) = \mu_2 \mathbf{u}_2$ podemos escribir

$$Q(x, y) = Q(s\mathbf{u}_1 + t\mathbf{u}_2) = \langle s\mathbf{u}_1 + t\mathbf{u}_2 \mid sL(\mathbf{u}_1) + tL(\mathbf{u}_2) \rangle =$$

$$= \langle\ s\mathbf{u}_1 + t\mathbf{u}_2 \mid \mu_1 s\mathbf{u}_1 + \mu_2 t\mathbf{u}_2)\ \rangle = \mu_1 s^2 + \mu_2 t^2.$$

luego $(x, y) = s\mathbf{u}_1 + t\mathbf{u}_2$ pertenece a N_c si y sólo sí $\mu_1 s^2 + \mu_2 t^2 = c$.

Si $AC - B^2 > 0$ y $A > 0$ sabemos que $0 < \mu_1 < \mu_2$, y la ecuación de N_c se puede escribir en la forma

$$a^2 s^2 + b^2 t^2 = c \quad \text{con} \quad a = \sqrt{\mu_1}, \quad b = \sqrt{\mu_2}$$

luego sólo hay curvas de nivel para $c > 0$ y estas son elipses que van aumentando de tamaño conforme $c > 0$ crece. La función Q toma el valor c en algún punto de M si y sólo si $M \cap N_c \neq \emptyset$, luego el máximo de Q sobre M será el mayor valor de c para el cual podamos asegurar que $M \cap N_c \neq \emptyset$. Esto ocurre precisamente cuando la elipse N_c es tangente, por fuera, a la circunferencia M; los dos puntos de tangencia diametralmente opuestos, $\mathbf{u}_1, -\mathbf{u}_1 \in M$, donde se alcanza el máximo $m_1 = \max Q(M)$ determinan la dirección común del eje mayor de las elipses. Análogamente el mínimo valor de Q sobre M es el valor de c para el cual la elipse N_c es tangente, por dentro, a la circunferencia M. El mínimo $m_2 = \min Q(M)$ se alcanza en dos puntos de tangencia diametralmente opuestos $\mathbf{u}_2, -\mathbf{u}_2$, que determinan la dirección común del eje menor de las elipses. Evidentemente $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales.

Cuando $AC - B^2 > 0$ y $A < 0$ la interpretación geométrica se reduce a la acabamos de hacer, considerando la función $-Q$.

Si $AC - B^2 < 0$ sabemos que $\mu_1 < 0 < \mu_2$, y la ecuación de N_c se puede escribir en la forma

$$a^2 s^2 - b^2 t^2 = c \quad \text{con} \quad a = \sqrt{-\mu_1}, \quad b = \sqrt{\mu_2}$$

Ahora las curvas de nivel N_c son hipérbolas que llenan el abierto

$$A = \{(s, t) : (as + bt)(as - bt) \neq 0\}.$$

Si $c > 0$ las hipérbolas ocupan $A^+ = \{(s, t) : (as+bt)(as-bt) > 0\}$, formado por dos de las regiones angulares opuestas que determinan las rectas $as+bt = 0$, $as-bt = 0$. Cuando $c < 0$ las hipérbolas ocupan, $A^- = \{(s, t) : (as + bt)(as - bt) < 0\}$, formado por las otras dos regiones angulares limitadas por las mismas rectas. Estas rectas son las asíntotas comunes de todas las hipérbolas, y conforme c se aproxima a 0 las hipérbolas van quedando más cerca de las asíntotas.

Ahora el máximo de Q sobre M es positivo. Es el mayor valor de c para el cual $M \cap N_c \neq \emptyset$. Esto ocurre precisamente para la única hipérbola N_c, con $c > 0$, que es tangente a la circunferencia M. Hay dos puntos de tangencia, diametralmente opuestos, $\mathbf{u}_1, -\mathbf{u}_1 \in M$, donde Q alcanza el máximo $m_1 = \max Q(M)$. Análogamente el mínimo de Q sobre M es negativo. Es el único $c < 0$ que hace que la hipérbola N_c sea tangente a la circunferencia M. Hay dos puntos de tangencia diametralmente opuestos, $\mathbf{u}_2, -\mathbf{u}_2 \in M$, que están en una recta perpendicular a la determinada por $\mathbf{u}_1$ y $-\mathbf{u}_1$. ■

Ejercicio 9.18 *Calcule los extremos absolutos de la forma cuadrática* $Q(x, y, z) = x^2 + y^2 + z^2 + 4xy$ *sobre la circunferencia*

$$C = \{(x, y, z) : x^2 + y^2 + z^2 = 1, \quad x + y + 2z = 0\}$$

Estudie los planos de la forma $z = ax - ay$ sobre los que la forma cuadrática es definida positiva.

SOLUCIÓN

Los extremos absolutos de Q sobre C se pueden obtener de varias formas
a) En la forma usual, como un problema de extremos con dos ligaduras: C es compacto (cerrado y acotado) luego la función continua Q alcanza sobre C un mínimo absoluto μ_1, y un máximo absoluto μ_2, es decir, existen $\mathbf{p}_1, \mathbf{p}_2 \in C$ tales que

$$\mu_1 = \text{mín}\{Q(\mathbf{p}) : \mathbf{p} \in C\} = Q(\mathbf{p}_1); \quad \mu_2 = \text{máx}\{Q(\mathbf{p}) : \mathbf{p} \in C\} = Q(\mathbf{p}_2);$$

Se comprueba fácilmente que en todo punto $(x, y, z) \in C$ los vectores $(2x, 2y, 2z)$, $(1, 1, 2)$ son linealmente independientes, luego C es una subvariedad de dimensión 1 y clase C^∞. Se sigue que $\mathbf{p}_1$ y $\mathbf{p}_2$ son puntos estacionarios de Q sobre C y según el método de los multiplicadores de Lagrange, estos puntos son soluciones del sistema

$$\begin{aligned} i)&\quad 2x + 4y - 2\mu x - \lambda = 0 \\ ii)&\quad 2y + 4x - 2\mu y - \lambda = 0 \\ iii)&\quad 2z - 2\mu z - 2\lambda = 0 \\ iv)&\quad x^2 + y^2 + z^2 = 1 \\ v)&\quad x + y - 2z = 0 \end{aligned}$$

Multiplicando la primera ecuación por y, la segunda por x y sumando miembro a miembro resulta $4(y^2 - x^2) = \lambda(y - x)$ de donde se obtiene que, o bien $(y - x) = 0$ o bien $4(y + x) = \lambda$. Con la primera alternativa, y utilizando las ecuaciones iv) y v) se obtienen los puntos estacionarios

$$\mathbf{a} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), \quad -\mathbf{a} = (-1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3})$$

Con la segunda alternativa, sustituyendo $\lambda = 4(x + y)$ en i) y ii), se obtiene $x = -y$, y utilizando otra vez iv) y v) salen los puntos estacionarios

$$\mathbf{b} = (1/\sqrt{3}, -1/\sqrt{3}, 0), \quad -\mathbf{b} = (-1/\sqrt{3}, 1/\sqrt{3}, 0)$$

Se concluye así que el máximo absoluto es $4/3 = Q(\mathbf{a}) = Q(-\mathbf{a})$ y que el mínimo absoluto es $-1 = Q(\mathbf{b}) = Q(-\mathbf{b})$.
b) También podemos empezar como en a) y terminar como en la demostración de la proposición I.2 (estamos considerando un caso particular de la situación considerada allí). Ahora, en vez de resolver completamente el sistema que proporciona los puntos estacionarios, podemos razonar como en I.2 y obtener que si μ, λ son los multiplicadores asociados a un punto estacionario $\mathbf{p}$ entonces $\mu = Q(\mathbf{p})$. Se sigue de esto que el mínimo y el máximo de Q sobre C vienen dados por las soluciones $\mu_1 = -1$, $\mu_2 = 7/5$ de la ecuación de segundo grado

$$\begin{vmatrix} (1-\mu) & 2 & 0 & 1 \\ 2 & (1-\mu) & 0 & 1 \\ 0 & 0 & (1-\mu) & 2 \\ 1 & 1 & 2 & 0 \end{vmatrix} = 0$$

c) Otra alternativa para resolver la primera parte del problema consiste en reducirlo a uno de extremos condicionados de una función de dos variables con una sola ligadura: Para $(x, y, z) \in C$ el valor $Q(x, y, z) = 1 + 4xy$ no depende de z, luego el problema es equivalente al de obtener los extremos absolutos de la función de dos variables $1 + 4xy$ sobre la elipse $x^2 + y^2 + (x + y)^2/4 = 1$ (la proyección de C sobre el plano (x, y)). Probablemente este es el camino más breve.

Para discutir cuando la forma cuadrática es definida positiva sobre el plano $z = ax - ay$ proponemos dos alternativas:
i) Restringir $Q(x, y, z)$ al plano $z = ax - ay$. Se obtiene así una forma cuadrática en dos variables $q(x, y) = Ax^2 + Cy^2 + 2Bxy$ con $A = C = 1 + a^2$, $B = 2 - a^2$, que es definida positiva cuando $A > 0$ y $AC - B^2 > 0$, lo que ocurre si y sólo si $a^2 > 1/2$.
ii) Acudir otra vez a la proposición I.2, que nos dice que la forma cuadrática es Q es definida positiva sobre el plano $z = ax - ay$ cuando son positivas las dos soluciones de la ecuación de segundo grado

$$\begin{vmatrix} (1-\mu) & 2 & 0 & a \\ 2 & (1-\mu) & 0 & -a \\ 0 & 0 & (1-\mu) & -1 \\ a & -a & -11 & 0 \end{vmatrix} = 0$$

Resolviendo la ecuación se obtiene fácilmente que ambas soluciones son positivas si y sólo si $a^2 > 1/2$. ■

9.4. Ejercicios propuestos

◇ **9.4.1** *Utilice el método de los multiplicadores de Lagrange para obtener, en cada caso, los extremos absolutos y relativos de la función* $f(x,y)$ *sobre la elipse* E*:*

a)	$f(x,y) = x + y^2$;	$E = \{(x,y) : 2x^2 + y^2 = 1\}$.
b)	$f(x,y) = x^2 + y^2 - 4xy + 20x + 20y$;	$E = \{(x,y) : x^2 + y^2 + xy = 12\}$.
c)	$f(x,y) = xy - 4x - 4y$;	$E = \{(x,y) : x^2 + y^2 + xy = 12\}$.

Interprete los resultados considerando las curvas de nivel $N_c = \{(x,y) : f(x,y) = c\}$.

◇ **9.4.2** *Utilice el método de los multiplicadores de Lagrange para obtener, en cada caso, los extremos absolutos de* $f(x,y)$ *sobre el compacto* $K \subset \mathbb{R}^2$*:*

a)	$f(x,y) = x - x^2 - y^2$;	$K = \{(x,y) : x \geq 0, x^2 + y^2 \leq 1\}$.
b)	$f(x,y) = \operatorname{sen}(xy)$;	$K = \{(x,y) : x^2 + y^2 \leq 1\}$.
c)	$f(x,y) = x^2 + y^2 + xy + x$	$K = \{(x,y) : x^2 + y^2 \leq 1\}$.

◇ **9.4.3** *Utilice el método de los multiplicadores de Lagrange para obtener:*
a) La mínima distancia entre la recta $y - x + 5 = 0$ *y la parábola* $y = x^2$.
b) La máxima y mínima distancia del origen a la elipse $5x^2 + 6xy + 5y^2 = 8$.

◇ **9.4.4** *Si* $g : \mathbb{R}^3 \to \mathbb{R}$ *es continua y* $S = \{(x,y,z) : g(x,y,z) = 0\}$ *no es vacío demuestre para cada* $\mathbf{p} \in \mathbb{R}^3 \setminus S$ *hay un punto* $\mathbf{q} \in S$ *cuya distancia a* $\mathbf{p}$ *es mínima:*

$$\|\mathbf{p} - \mathbf{q}\|_2 = \text{mín}\{\|\mathbf{p} - \mathbf{x}\|_2 : \mathbf{x} \in S\}$$

Si g *es de clase* C^1 *y* $\nabla g(x,y,z) \neq 0$ *para todo* $(x,y,z) \in S$ *justifique que el vector* $\mathbf{p} - \mathbf{q}$ *es normal a la superficie* S *en el punto* $\mathbf{q}$. *Calcule la mínima distancia de* $\mathbf{p} = (1,1,1)$ *a la superficie* $x^2 + y^2 - z^2 - 2x + 2 = 0$.

◇ **9.4.5** *Utilice el método de los multiplicadores de Lagrange para obtener los puntos de la superficie* $z^2 - xy - 1 = 0$ *más próximos al origen.*

◇ **9.4.6** *Obtenga los puntos de la curva* $x^2 - xy + y^2 - z^2 = 1$, $x^2 + y^2 = 1$, *que están más cerca del origen.*

◇ **9.4.7** *Obtenga los puntos de* $C = \{(x,y,z) : 2z + x^2 + y^2 = 16; x + y = 4\}$ *que están mas cerca del origen. Compruebe que* $C_1 = \{(x,y,z) \in C : x \geq 0, y \geq 0, z \geq 0\}$ *es compacto y obtenga los puntos de* C_1 *que están más lejos del origen.*

◇ **9.4.8** *Sean* $S, T \subset \mathbb{R}^3$ *subvariedades diferenciables disjuntas de clase* C^1, *tales que existen* $\mathbf{p} \in S$, $\mathbf{q} \in T$ *verificando* $\|\mathbf{p} - \mathbf{q}\|_2 = \text{mín}\{\|\mathbf{x} - \mathbf{y}\|_2 : \mathbf{x} \in S, \mathbf{y} \in T\}$. *Demuestre que el vector* $\mathbf{p} - \mathbf{q}$ *es ortogonal a* S *y* T *en* $\mathbf{p}$ *y* $\mathbf{q}$ *respectivamente.*

◇ **9.4.9** *Utilice el método de los multiplicadores de Lagrange para obtener la mínima distancia entre la superficie* $z = x^2 + y^2$ *y el plano* $x + 2y - z = 4$.

♢ **9.4.10** *Se supone que el plano* $P = \{(x, y, z) : Ax + By + Cz = D\}$ *no corta a la superficie* $S = \{(x, y, z) : x^2 + y^2 - z = 0\}$.

Justifique que la distancia entre un punto genérico $\mathbf{p} \in P$ *y un punto genérico* $\mathbf{q} \in S$ *alcanza un mínimo absoluto en un único par de puntos* $\mathbf{a} \in P$,$\mathbf{b} \in S$. *Calcule* $\mathbf{a}$ *y un vector normal a* S *en* $\mathbf{b}$. *Obtenga la mínima distancia entre* S *y* $P = \{(x, y, z) : x + 2y - z = 4\}$.

♢ **9.4.11** *Calcule la máxima y la mínima distancia del origen a la elipse*

$$E_a = \{(x, y, z) : x^2 + y^2 \leq 1, x + y + az = 1\}$$

♢ **9.4.12** *De todos los paralelepípedos inscritos en el elipsoide* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$, *obtenga el de mayor volumen. Calcule también el mínimo volumen encerrado por un plano tangente al elipsoide y los planos* $x = 0$, $y = 0$, $z = 0$.

♢ **9.4.13** *De todos los planos tangentes al trozo de elipsoide*

$$S = \{(x, y, z) : x^2 + 2y^2 + 3z^2 = 1,\ x > 0\ y > 0\ z > 0\}$$

determine el que forma con los planos $x = 0$, $y = 0$, $z = 0$ *un tetraedro de volumen mínimo.*

♢ **9.4.14** *Determine el elipsoide* $E(a, b, c) = \{(x, y, z) : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}$ *que pasa por* $\mathbf{p} = (1, 2, 3)$ *y encierra mayor volumen.*

♢ **9.4.15** *Utilice el método de los multiplicadores de Lagrange para determinar las dimensiones de una caja rectangular sin tapa, con superficie de* 16 m^2, *que encierra un volumen máximo. Justifique la existencia de la caja de volumen máximo.*

♢ **9.4.16** *Determine las dimensiones del bote cilíndrico (con tapa) de mayor superficie que se puede inscribir en una esfera de radio* R.

♢ **9.4.17** *Obtenga los extremos relativos de la función* $3x-4y+2z$ *sobre la superficie* $x^3 - 2y^2 + z^2 = 0$.

♢ **9.4.18** *Obtenga los extremos absolutos y relativos de las siguientes funciones sobre las superficies que se indican:*

a)	$x - 2y + 2z$,	*sobre la esfera*	$x^2 + y^2 + z^2 = 1$.
b)	$x + y + z$,	*sobre el elipsoide*	$x^2 + 2y^2 + 3z^2 = 1$.
c)	$x + y + z$,	*sobre el hiperboloide*	$x^2 + y^2 - 2z^2 = 6$.
d)	$x^2 + y^2 + z^2$,	*sobre la superficie*	$x^2 + y^2 - z^2 + 2xy = 16$.
e)	$x^2 + y^2 + 8xy + 10z$	*sobre la esfera*	$x^2 + y^2 + z^2 = 3$.

♢ **9.4.19** *Estudie si la función* $x^2 + y^2 + z^2$ *posee en el punto* $(0, 1, 0)$ *un extremo condicionado por la ligadura* $z^2 + 2xyz + y^2 + x^3 = 1$

♢ **9.4.20** *Calcule el máximo y el mínimo absoluto de la función* x^2+y^2+z *sobre la curva* $C = \{(x,y,z) : x^2+y^2+z^2 = 4, x^2+y^2 = 2x\}$*. Compruebe que existe* $\mathbf{p} \in C$ *tal que* $C_0 = C \setminus \{\mathbf{p}\}$ *es subvariedad diferenciable de* $\mathbb{R}^3$*.*

♢ **9.4.21** *a) Compruebe que* $C = \{(x,y,z) \in \mathbb{R}^3 : 3z^2 - 2xy - 1 = 0, x^2+y^2-2z^2 = 0\}$ *es una subvariedad diferenciable de clase* C^∞ *y dimension* 1*.*
b) Justifique que el punto $\mathbf{p} = (1,1,1)$ *posee un entorno abierto* $U_p \subset \mathbb{R}^3$ *tal que el trozo de curva* $C \cap U_p$ *admite una parametrización de clase* C^∞ *la forma* $\gamma(t) = (t, y(t), z(t))$*. Calcule los vectores* $\mathbf{u} = \gamma'(1)$ *y* $\mathbf{v} = \gamma''(1)$*.*
c) Sea $f(x,y,z) = x + y^3 - 2y + z^2$*. Calcule* $D_{\mathbf{u}} f(\mathbf{p})$*. Compruebe que* $f|_C$ *presenta un extremo relativo en* $\mathbf{p}$ *y determine su naturaleza.*

♢ **9.4.22** *Compruebe que* $C = \{(x,y,z) \in \mathbb{R}^3 : x+y = 1,\ x^2+y^2+z^2 = 1\}$ *es una subvariedad diferenciable de* $\mathbb{R}^3$ *de clase* C^∞ *y dimensión* 1*, y que* $\mathbf{p} = (0,1,0)$ *es un punto estacionario de* $f(x,y,z) = x(y+z)$ *sobre la curva* C*. Estudie si* $f|_C$ *presenta en* $\mathbf{p}$ *un extremo relativo.*

♢ **9.4.23** *Compruebe que la curva* $C = \{(x,y,z) : x^2 - y^2 - 1 = 0, 2x + z - 1 = 0\}$ *es una subvariedad diferenciable de clase* C^∞ *y dimensión* 1*, y escriba la ecuación de la recta tangente a* C *en un punto genérico* $(x_0, y_0.y_0) \in C$*.*
Calcule los puntos estacionarios de $f(x,y,z) = x + y^2 + z$ *sobre* C *y compruebe que* f *presenta mínimos relativos en todos ellos.*
Estudie si alguno de los mínimos relativos es un mínimo absoluto.

♢ **9.4.24** *Compruebe que la curva* $C = \{(x,y,z) : x^2+y^2 = z^2,\ y = \sqrt{3}x + 2\}$ *es una subvariedad diferenciable de clase* C^∞ *y dimensión* 1*. Determine los extremos relativos y absolutos de la función* $f(x,y,z) = x^2+y^2+(z-1)^2$ *sobre la curva. Interpretación geométrica del resultado.*

♢ **9.4.25** *Compruebe que* $\mathbf{p} = (\sqrt{2}/4, -1/2, 0)$*,* $\mathbf{q} = (\sqrt{2}/4, 1/2, 0)$ *son puntos estacionarios de la función* $f(x,y,z) = 2x^2+y^2+z^2-xy$ *sobre el elipsoide* $S = \{(x,y,z) : 4x^2+2y^2+z^2 = 1\}$*, y estudie su naturaleza.*
¿Hay más puntos estacionarios?. Justifique la respuesta.

♢ **9.4.26** *Compruebe que* $\mathbf{p} = (1/3, -1/3, -5/3)$ *y* $\mathbf{q} = (1,1,1)$ *son puntos estacionarios de la función* $x^2+y^2+8xy+10z$ *sobre la esfera* $x^2+y^2+z^2 = 3$*. Determine su naturaleza.*

♢ **9.4.27** *Calcule los extremos relativos y absolutos de* $f(x,y,z) = z - (1+x^2+y^2)e^z$ *sobre la elipse* $E = \{(x,y,z) : x^2+y^2 = 1, x+z = 1\}$*. Si* $\mathbf{p} \in E$ *es el punto donde* $f|_E$ *alcanza el mínimo absoluto, obtenga la ecuación de la recta tangente a* E *en* p*.*

♢ **9.4.28** *Compruebe que* $C = \{(x,y,z) : x^2+y^2-2 = 0,\ x - yz + 1 = 0\}$ *es una subvariedad de dimensión* 1 *y clase* C^∞*.*
a) Obtenga un vector $\mathbf{w} = (w_1, w_2, w_3)$ *tangente a* C *en* $\mathbf{p} = (1,1,2)$*.*
b) Para la función $f(x,y,z) = xy - ayz + z$ *calcule la derivada* $D_{\mathbf{w}} f(\mathbf{p})$*.*
c) Determine $a \in \mathbb{R}$ *para que* $\mathbf{p}$ *sea punto estacionario de* $f|_C$ *y en ese caso estudie su naturaleza (e.d. si es punto máximo o mínimo relativo).*

♢ **9.4.29** *En cada caso calcule los extremos absolutos sobre K, de la función dada:*

a) $x^2+y^2+z^2+x+y+z$; $K=\{x^2+y^2+z^2\leq 4, z\leq 1\}$.

b) $x+y+z$; $K=\{0\leq z\leq 2, x^2+y^2-2z^2\leq 6\}$.

c) $x+y-\sqrt{2}z+x^2+y^2+z^2$; $K=\{0\leq z\leq 1, x^2+y^2=2z^2\}$.

d) $x^2+y^2+z^2+xy+x-2z$; $K=\{x^2+y^2\leq 1, |z|\leq 2\}\}$.

e) $x^2-y^2+z^2-z$; $K=\{x^2+y^2+z^2\leq 1, x^2+y^2\geq z^2, z\geq 0\}$.

f) $x+y+z$; $K=\{x^2+y^2\leq z\leq 2\}$.

g) $x^2+y^2+z^2+x+y-\sqrt{2}z$; $K=\{x^2+y^2-z^2=0;\quad 0\leq z\leq 1\}$

♢ **9.4.30** *Sea $(a_{ij})_{1\leq i,j\leq n}$ una matríz simétrica, y $A:\mathbb{R}^n\to\mathbb{R}^n$, la aplicación lineal asociada $A(\mathbf{x})_i=\sum_{j=1}^n a_{ij}x_j$.*

i) Demuestre que el máximo valor μ del polinomio $p(\mathbf{x})=\sum_{i,j=1}^n a_{ij}x_ix_j$ sobre la esfera $\{(x_1,x_2,\cdots x_n): x_1^2+x_2^2+\cdots x_n^2=1\}$ es el máximo autovalor de A.

ii) En el caso $n=2$ se considera un polinomio $p(x,y)=ax^2+2bxy+cy^2$ tal que $C=\{(x,y)\in\mathbb{R}^2: p(x,y)=1\}$ no es vacío. Demuestre que $\mu>0$ y que la minima distancia de la cónica C al origen es $1/\sqrt{\mu}$. Si C es una elipse, ¿Cuál es la máxima distancia de C al origen?.

♢ **9.4.31** *Sea $(a_{ij})_{1\leq i,j\leq 3}$ una matriz simétrica. Demuestre que los extremos de $f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2$ sujetos a las ligaduras*

$$\sum_{1\leq i,j\leq 3} a_{ij}x_ix_j=1;\quad a_1x_1+a_2x_2+a_3x_3=0$$

son $1/\mu_1$ y $1/\mu_2$ donde μ_1,μ_2 son las soluciones de la ecuación de segundo grado

$$\begin{vmatrix} a_1 & a_2 & a_3 & 0\\ a_{11}-\mu & a_{12} & a_{13} & a_1\\ a_{21} & a_{22}-\mu & a_{23} & a_2\\ a_{31} & a_{32} & a_{33}-\mu & a_3\end{vmatrix}=0$$

Mediante una interpretación geométrica del resultado justifique que esta ecuación ha de tener dos raíces reales positivas.