

Variedades. Extremos condicionados.

Los resultados que hemos visto hasta ahora permiten localizar los puntos donde un campo escalar diferenciable alcanza un máximo o mínimos relativos, pero no se pueden usar cuando el conjunto en el que está definida la función cuyo máximo ó mínimo pretendemos calcular tiene interior vacío. Este es el caso, por ejemplo, de que el dominio sea una recta R en \mathbb{R}^2 y se pretenda calcular el máximo de una función sobre ésta recta (si existe). En este caso, la idea clave es considerar la recta como un espacio topológico que es homeomorfo a \mathbb{R} . Si $p : \mathbb{R} \rightarrow R$ es un homeomorfismo que sea derivable (los hay afines) y $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ es la función cuyo máximo o mínimo queremos averiguar sobre la recta, suponiendo que f sea diferenciable, y $c \in \mathbb{R}$ es tal que $p(c) = a \in R$, entonces es claro que

$$f(x) \leq f(a), \forall x \in R \Leftrightarrow f(p(t)) \leq f(p(c)) \quad \forall t \in \mathbb{R}.$$

En caso de que se verifique la condición anterior, la función $h = f \circ p$ alcanza el máximo en c . Como h es derivable, en virtud de la Regla de la cadena, entonces ha de verificarse que $h'(c) = 0$, esto es, $Df(a)(p'(c)) = 0$. Esta idea tan simple se puede aplicar si en lugar de ser \mathbb{R} el dominio de p , p es un homeomorfismo diferenciable de un abierto U de \mathbb{R}^k sobre un subconjunto de \mathbb{R}^n y en este caso f es un campo escalar diferenciable definido en un abierto de \mathbb{R}^n que contiene a $p(U)$. Por supuesto, en tal caso aplicando la Regla de la cadena se obtiene que

$$Df(a) \circ Dp(c) = 0, \quad \text{esto es,} \quad Jf(a)Jp(c) = 0,$$

si $f \circ p$ tiene un máximo relativo en c , donde $c \in U$ es tal que $p(c) = a$.

Una recta en \mathbb{R}^2 es un caso muy particular del concepto de variedad que presentamos a continuación.

10.1 Definición. Sea $M \subset \mathbb{R}^n$ y $1 \leq k < n$. Diremos que M es una variedad (diferenciable de clase C^1) de dimensión k si para cada punto $a \in A$ existe un abierto U de \mathbb{R}^k y una aplicación $p : U \rightarrow \mathbb{R}^n$ tal que p es de clase $C^1(U)$, el rango de $Jp(x)$ es k para cada $x \in U$ y además p es un homeomorfismo de U sobre $M \cap W$, donde W es un abierto de \mathbb{R}^n que contiene a a . En tal caso diremos que p es una parametrización local de p en a .

Aunque la definición anterior no será la forma habitual en que describiremos las variedades, consideramos el siguiente ejemplo trivial de variedad.

Si $a, b, c \in \mathbb{R}$ son tales que $a^2 + b^2 \neq 0$ y consideramos la recta dada por

$$R = \{(x, y) \in \mathbb{R}^2 : ax + by = c\},$$

entonces R es una variedad de dimensión 1. Por hipótesis, alguno de los números a ó b es no nulo. Si suponemos que b es distinto de cero, basta considerar la aplicación $p : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por

$$p(x) = \left(x, \frac{c - ax}{b}\right).$$

Es inmediato comprobar que la aplicación p verifica las condiciones de la definición anterior. En caso de que $a \neq 0$, basta definir

$$p(y) = \left(\frac{c - by}{a}, y\right) \quad (y \in \mathbb{R}).$$

El siguiente resultado caracteriza las variedades. De hecho, habitualmente las variedades aparecen como conjuntos de ceros de una función, y corresponden a la descripción dada en el tercer apartado del resultado que enunciamos a continuación.

10.2 Proposición. Sea $M \subset \mathbb{R}^n$ y $1 \leq k < n$. Las siguientes afirmaciones son equivalentes:

i) M es una variedad de dimensión k .

ii) Para cada punto $a \in M$, salvo reordenación de las variables, existe un abierto $U \subset \mathbb{R}^k$, un abierto $W \subset \mathbb{R}^n$ que contiene a a y una aplicación $p : U \rightarrow \mathbb{R}^{n-k}$ de clase $C^1(U)$ tal que

$$M \cap W = \{(x, p(x)) : x \in U\}.$$

iii) Para cada punto $a \in M$, salvo reordenación de las variables, existe un abierto $G_a \subset \mathbb{R}^n$, un abierto $W \subset \mathbb{R}^n$ que contiene a a y una aplicación $F : G_a \rightarrow \mathbb{R}^{n-k}$ de clase $C^1(G_a)$ tal que

$$M \cap W = \{x \in G_a : F(x) = 0\}$$

y además el rango de $JF(x) = n - k$ para cada $x \in G_a$.

Para la prueba de i) \Rightarrow ii) se usa el teorema de la función inversa aplicado a la composición de la parametrización local con una conveniente proyección de \mathbb{R}^n sobre \mathbb{R}^k . Por ejemplo, si las k últimas columnas de $Jp(c)$ dan una matriz inversible, entonces se usa la proyección sobre las k últimas coordenadas en \mathbb{R}^n .

Las implicaciones ii) \Rightarrow i) y ii) \Rightarrow iii) son inmediatas y la implicación iii) \Rightarrow ii) es consecuencia del Teorema de la función implícita.

En caso de que se verifique el apartado c) del resultado anterior, diremos que F determina implícitamente la variedad M en un entorno de a .

El subespacio tangente a una variedad en un punto que definimos a continuación juega un papel esencial en el estudio de los extremos de una función sobre una variedad.

10.3 Definición. Sea $M \subset \mathbb{R}^n$ una variedad de dimensión k . El subespacio (vectorial) tangente a M en un punto $a \in M$, que notaremos por $T_M(a)$, viene dado por

$$T_M(a) = \{\gamma'(0) : \gamma :]-\delta, \delta[\rightarrow M, \gamma \text{ es continua, derivable en } 0 \text{ y } \gamma(0) = a\}.$$

Notaremos por $T_M(a)^\perp$ al subespacio ortogonal a $T_M(a)$, que llamaremos subespacio normal a M en a , esto es,

$$T_M(a)^\perp = \{y \in \mathbb{R}^n : \langle y, x \rangle = 0, \forall x \in T_M(a)\}.$$

Aunque en el estudio de extremos de un campo escalar condicionados por una variedad los dos subespacios anteriores juegan un papel esencial, la variedad afín tangente a una variedad M en un punto $a \in M$ viene dada por

$$a + T_M(a)$$

y la variedad afín normal a M en un punto $a \in M$ por

$$a + T_M(a)^\perp.$$

El siguiente resultado permite en casos prácticos describir fácilmente el espacio tangente a la variedad en un punto.

10.4 Proposición. Sea $M \subset \mathbb{R}^n$ una variedad de dimensión k y $a \in M$, entonces se tiene que

$$T_M(a) = Dp(c)(\mathbb{R}^k) = \text{Ker } DF(a),$$

donde p es una parametrización local de M en a , $p(c) = a$ y F es una aplicación que define a M en un entorno de a .

Como consecuencia, $T_M(a)$ es un subespacio vectorial dimensión k y $T_M(a)^\perp$ es el subespacio vectorial de dimensión $n - k$ generado por los vectores

$$\{\nabla F_i(a) : 1 \leq i \leq n - k\}.$$

Para probar el resultado anterior se comprueba que

$$Dp(c)(\mathbb{R}^k) \subset T_M(a) \subset \text{Ker } DF(a).$$

La primera inclusión se obtiene sin más que considerar la curva $\gamma(t) = p(c + tv)$ donde $v \in \mathbb{R}^k$, definida en un intervalo abierto que contiene a 0. Para la segunda inclusión basta usar que $F(\gamma(t)) = 0$ para cualquier curva γ de las que aparece en la definición de $T_M(a)$. Basta usar después la Regla de la cadena.

Por las condiciones sobre el rango del jacobiano de p y de F en cada punto, se obtiene que los subespacios vectoriales $Dp(c)(\mathbb{R}^k)$ y $\text{Ker } DF(a)$ tienen dimensión k , luego coinciden entre sí, y también coinciden con $T_M(a)$. Del argumento anterior se obtiene también la descripción de $T_M(a)^\perp$.

La siguiente noción juega un papel esencial en este tema.

10.5 Definición. Sea $M \subset \mathbb{R}^n$ una variedad, $a \in M$ y $f : G \rightarrow \mathbb{R}$ una función definida en un abierto $G \subset \mathbb{R}^n$ que contiene al punto a . Diremos que la función f tiene un máximo local en a condicionado por M si existe $r > 0$ tal que

$$B(a, r) \subset G \text{ y } f(x) \leq f(a), \forall x \in M \cap B(a, r).$$

El campo escalar f tiene un mínimo local en a condicionado por M si existe $r > 0$ tal que

$$B(a, r) \subset G \text{ y } f(x) \geq f(a), \forall x \in M \cap B(a, r).$$

En cualquiera de los dos casos anteriores, diremos que f tiene un extremo local en a condicionado por M .

Como la parametrización local p de M en a es un homeomorfismo de un abierto de \mathbb{R}^k sobre un entorno de a en la topología relativa de M tenemos:

10.6 Proposición. Sea $M \subset \mathbb{R}^n$ una variedad, $a \in M$ y $f : G \rightarrow \mathbb{R}$ una función definida en un abierto $G \subset \mathbb{R}^n$ que contiene a a . Se verifica que entonces que f tiene un extremo local en a condicionado por M si, y sólo si, $f \circ p$ tiene un extremo relativo en c , donde $p(c) = a$.

Además, los extremos de f y de $f \circ p$ son del mismo tipo (ambos son máximos o ambos mínimos).

La equivalencia anterior y los resultados conocidos que relacionan los extremos relativos de un campo escalar diferenciable con los puntos críticos son la clave del resultado que permite localizar los posibles extremos de un campo escalar condicionados por una variedad. Antes de enunciar el resultado correspondiente presentamos la siguiente notación.

Sea $f : W \subset \mathbb{R}^n \rightarrow \mathbb{R}$ un campo escalar definido en un abierto W que corta a una variedad M de dimensión k y supongamos que F define a M en un entorno de un punto $a \in M \cap W$. La función de Lagrange G asociada al campo escalar f y la variedad M está dada por

$$G(x, \mu) = f(x) + \sum_{i=1}^{n-k} \mu_i F_i(x) \quad (x \in W, \mu \in \mathbb{R}^{n-k}).$$

10.7 Teorema. Sea $M \subset \mathbb{R}^n$ una variedad de dimensión k determinada por F en un punto $a \in M$ y $f : W \rightarrow \mathbb{R}$ un campo escalar diferenciable definido en un abierto $W \subset \mathbb{R}^n$ que contiene a a . Si f tiene en a un extremo local condicionado por la variedad, existe un único $\mu \in \mathbb{R}^{n-k}$ tal que el par (a, μ) es un punto crítico de la función de Lagrange G .

Idea de la prueba: Si f tiene en a un extremo local condicionado por M , por la Proposición 10.6, entonces $f \circ p$ tiene un extremo relativo en c , donde p es una parametrización local de M en a y $p(c) = a$. Como f y p son diferenciables, por la regla de la cadena el campo escalar $f \circ p$ también lo es, luego sabemos que $D(f \circ p)(c) = 0$, esto es,

$$Df(a) \circ Dp(c) = 0.$$

Como la imagen de $Dp(c)$ de \mathbb{R}^k es $T_M(a)$ en vista de la Proposición 10.4, el hecho de que $Df(a) \circ Dp(c) = 0$ significa que $T_M(a)$ está contenido en el núcleo de $Df(a)$, equivalentemente,

$$\langle \nabla f(a), x \rangle = 0, \quad \forall x \in T_M(a).$$

La igualdad anterior nos dice que $\nabla f(a) \in T_M(a)^\perp$, luego $\nabla f(a)$ es combinación lineal de los vectores $\{\nabla F_i(a) : 1 \leq i \leq n-k\}$, que forman una base del subespacio $T_M(a)^\perp$, en virtud de la Proposición 10.4.

Cuando se calcula el gradiente de G , se obtiene que

$$\frac{\partial G}{\partial \mu_i}(x, \mu) = F_i(x), \quad \forall 1 \leq i \leq n-k.$$

Luego la condición $\frac{\partial G}{\partial \mu_i}(a, \mu) = 0$ para $1 \leq i \leq n-k$ significa que $a \in M$, ya que F determina localmente a M . La condición

$$\frac{\partial G}{\partial x_i}(x, \mu) = 0, \quad \forall 1 \leq i \leq n$$

significa que

$$\nabla f(a) + \sum_{i=1}^{n-k} \mu_i \nabla F_i(a) = 0,$$

esto es, $\nabla f(a) \in T_M(a)^\perp$. Como consecuencia, $a \in M$ y además c es un punto crítico de $f \circ p$.

Igual que en el caso de extremos relativos hay un criterio que permite clasificar los puntos críticos usando el hessiano del campo escalar que se quiere optimizar, para extremos condicionados también existe un criterio usando el hessiano de la función de Lagrange. En este último caso, si (a, λ) es un punto crítico de G , se fija el último grupo de variables iguales a λ , se considera el hessiano de la función

$$x \mapsto G(x, \lambda)$$

y se restringe el hessiano a $T_M(a)$. Si la forma cuadrática asociada es definida positiva, se verifica que la forma cuadrática asociada a $f \circ p$ en c también lo es, luego $f \circ p$ tiene un mínimo relativo en c . Por tanto, f tiene un mínimo local en a condicionado por M . Análogamente, si la forma cuadrática asociada a la restricción del hessiano de la función anterior a $T_M(a)$ es definida negativa, entonces f tiene un máximo local en a condicionado por M .

Si quereis ver algún ejemplo de cómo se aplica el criterio anterior en algún caso concreto, podeis encontrarlo en el Tema 9 de los apuntes de Análisis Matemático II de Acosta, Aparicio, Moreno y Villena.

Por supuesto, en casos en que la variedad es compacta, por el Teorema de Weierstrass, sabemos que el campo escalar f (que suponemos continuo) ha de alcanzar el máximo y el mínimo absolutos en la variedad. Por tanto, los puntos donde se alcanzan el mínimo y el máximo han de ser soluciones del sistema de Lagrange si f es diferenciable, ya que en ellos f tiene extremos locales condicionadas por la variedad.