# Análisis Matemático I,

# 2º Doble Grado Informática-Matemáticas

## Capítulo IV: VARIEDADES DIFERENCIABLES EN $\mathbb{R}^N$

## Tema 12: VARIEDADES DIFERENCIABLES EN $\mathbb{R}^N$. EXTREMOS CONDICIONADOS

María D. Acosta

Universidad de Granada

6-12-2020

# Motivación

Si queremos optimizar un campo escalar en un conjunto con interior vacío, no es útil el método que hemos visto hasta ahora para estudiar extremos relativos. Por ejemplo, si el dominio del campo escalar es la esfera unidad de $\mathbb{R}^2$, dada por

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

En este caso, usaremos la idea de que "localmente" la esfera es homeomorfa a un intervalo de $\mathbb{R}$. Por eso empezaremos definiendo el concepto de variedad diferenciable (de dimensión 1 en este caso).

Antes de dar la definición y las caracterizaciones de las variedades, particularizamos para $S^1$ las condiciones de la definición y de las caracterizaciones:

**1)** Para cada punto $p \in S^1$, existe un abierto $U$ en $\mathbb{R}^2$ que contiene a $p$ y un intervalo abierto $I$ de $\mathbb{R}$ tal que $I$ y $U \cap S^1$ son homeomorfos. Además el homeomorfismo verifica condiciones "extra" de diferenciabilidad.
Por ejemplo, si $p = (1, 0)$, la aplicación
$g : ]-\pi, \pi[ \longrightarrow S^1 \cap \{(x, y) \in \mathbb{R}^2) : y = 0, x \leq 0\}$ dada por
$g(t) = (\cos(t), \mathrm{sen}(t))$ es un homeomorfismo.

# Motivación

**2)** $S^1$ es el conjunto de ceros de una función diferenciable cuya diferencial en cada punto de $S^1$ tiene rango máximo.
En este caso, basta tomar $f(x,y) = x^2 + y^2 - 1$. Como $\nabla f(x,y) = (2x, 2y)$, entonces el rango de $\nabla f(x,y)$ es 1 en puntos de la esfera.

**3)** En un entorno de cada punto de la variedad, $S^1$ viene dada por la gráfica de una función diferenciable definida en un intervalo de $\mathbb{R}$.
Por ejemplo, si $p = (0,1)$, la aplicación $p : ]-1,1[ \longrightarrow \mathbb{R}^2$ dada por $p(x) = (x, \sqrt{1-x^2})$ es diferenciable y su imagen es la mitad superior de $S^1$, es decir, la intersección de $S^1$ con $\{(x,y) \in \mathbb{R}^2 : y > 0\}$.

# Variedad diferenciable de $\mathbb{R}^N$

**Definición**

Sea $M \subset \mathbb{R}^N$. Se dice que $M$ es una **variedad diferenciable (de clase $C^q$) de dimensión** $k$ si para cada $a \in M$, existen

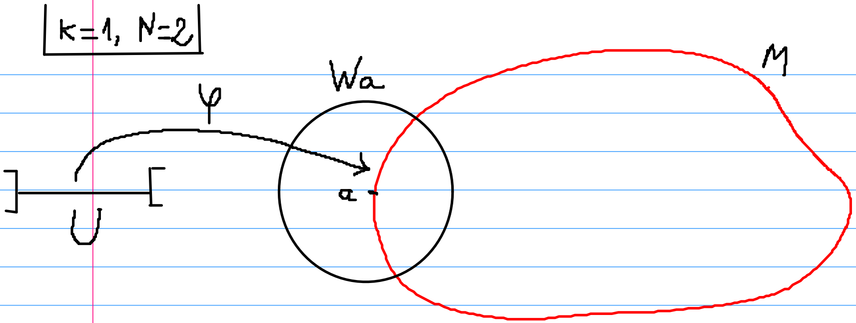
$$\begin{cases} W_a \text{ entorno abierto de } a \text{ en } \mathbb{R}^N \\ U \text{ abierto de } \mathbb{R}^k \\ p : U \longrightarrow \mathbb{R}^N \text{ homeomorfismo de clase } \mathcal{C}^q \text{ de } U \text{ sobre } p(U) \end{cases}$$

tales que $\begin{cases} p(U) = M \cap W_a \\ \operatorname{rango}\left(J_p(y)\right) = k, \quad \forall y \in U. \end{cases}$

La terna $(W_a, U, p)$ se denomina **una parametrización local de $M$ en $a$**.

# Variedad diferenciable de $\mathbb{R}^N$

**Idea intuitiva**



# Variedad diferenciable de $\mathbb{R}^N$

Las variedades vendrán dadas habitualmente como el conjunto de los ceros de un campo vectorial. Para trabajar con los ejemplos concretos y justificar varios resultados teóricos será útil la siguiente caracterización:

## Proposición

Sea $M \subset \mathbb{R}^N$, $1 \leq k < N$ y $q \in \mathbb{N}$. Equivalen las siguientes afirmaciones:

**1)** $M$ es una variedad diferenciable de dimensión $k$ y clase $C^q$.

**2)** Para cada punto $a \in M$ existe un abierto $U \subset \mathbb{R}^k$ y un abierto $G_a \subset \mathbb{R}^N$ tal que $a \in G_a$ y una aplicación $\varphi : U \longrightarrow \mathbb{R}^{N-k}$ de clase $C^q$ tal que

$$\{(x, \varphi(x)) : x \in U\} = G_a \cap M,$$

salvo una permutación de las coordenadas.

**3)** Para cada punto $a \in M$ existe un abierto $\Omega_a \subset \mathbb{R}^N$ tal que $a \in \Omega_a$ y una aplicación $F : \Omega \longrightarrow \mathbb{R}^{N-k}$ de clase $C^q$ tal que

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

y además el rango de $J_f(x) = N - k$ para cada punto $x \in \Omega_a$.

# Variedad diferenciable de $\mathbb{R}^N$

**Demostración:**
1) $\Rightarrow$ 2) Por hipótesis $M$ es una variedad diferenciable de dimensión $k$ y clase $C^q$. Luego, para cada $a \in M$ existen un abierto $W_a \subset \mathbb{R}^N$ con $a \in W_a$, un abierto $U \subset \mathbb{R}^k$ y una aplicación $p : U \longrightarrow \mathbb{R}^N$ de clase $C^q$ tal que $p$ es un homeomorfismo de $U$ en $W_a \cap M$ y además el rango de $Jp(x) = k$ para cada $x \in U$.
Supongamos que $u_0 \in U$ verifica que $p(u_0) = a$. Salvo una permutación de las variables en $\mathbb{R}^N$, podemos suponer que el rango del jacobiano de $Jp(u_0)$ viene dado por las $k$ primeras filas de $Jp(a)$. Sea $Q : \mathbb{R}^N \longrightarrow \mathbb{R}^k$ la proyección dada por

$$Q(y_1, \cdots, y_N) = (y_1, \cdots, y_k)$$

que es de clase $C^\infty$ por ser lineal. Dado que $g = Q \circ p = (p_1, \cdots, p_k)$, la aplicación $g$ es de clase $C^q$ por serlo $p$ y además $\det Jg(u_0) \neq 0$, ya que $Jg(x)$ es la matriz formada por las $k$ primeras filas de $Jp(x)$ para cada $x \in U$.
Podemos usar la versión local del Teorema de la función inversa. Luego existe un abierto contenido en $U$ que contiene a $u_0$ tal que $g$ es un difeomorfismo de clase $C^q$ de él en su imagen.

# Variedad diferenciable de $\mathbb{R}^N$

Cambiando $U$ por éste nuevo abierto, podemos suponer sin pérdida de generalidad que $g$ es un difeomorfismo de $U$ en su imagen $V \subset \mathbb{R}^k$.
Nótese que $g = (p_1, \cdots p_k)$.
Definimos $h : V \longrightarrow \mathbb{R}^N$ por

$$h(y)) = p(g^{-1}(y)), \qquad (y \in V).$$

Como $g$ es un difeomorfismo en $V$ de clase $C^q$, $h$ es de clase $C^q$. Si $y \in V$ existe un único $x \in U$ tal que $g(x) = y$. Es decir, $(p_1(x), \cdots, p_k(x)) = y$ luego

$$h(y) = (p_1(x), \cdots, p_k(x), p_{k+1}(g^{-1}(y)), p_{k+2}(g^{-1}(y)), \cdots, p_N(g^{-1}(y)) =$$

$$(y_1, \cdots, y_k, p_{k+1}(g^{-1}(y)), p_{k+2}(g^{-1}(y)), \cdots, p_N(g^{-1}(y)).$$

Como $g^{-1}(V) = U$ y $p(U)$ es un abierto relativo a la variedad que contiene al punto $a$, hemos comprobado que se verifica 2), donde

$$\varphi = (p_{k+1} \circ g^{-1}, p_{k+2} \circ g^{-1}, \cdots, p_N \circ g^{-1}).$$

# Variedad diferenciable de $\mathbb{R}^N$

2) $\Rightarrow$ 1) Por hipótesis, para cada punto $a \in M$ existe un abierto $U \subset \mathbb{R}^k$ y un abierto $G_a \subset \mathbb{R}^N$ tal que $a \in G_a$ y una aplicación $\varphi : U \longrightarrow \mathbb{R}^{N-k}$ de clase $C^q$ tal que

$$\{(x, \varphi(x)) : x \in U\} = G_a \cap M,$$

salvo una permutación de las coordenadas.

Consideramos la aplicación $p : U \longrightarrow \mathbb{R}^N$ dada por

$$p(x) = (x, \varphi(x)) \quad (x \in U),$$

que es biyectiva sobre su imagen y $C^q$ por serlo $\varphi$. Como la inversa sobre su imagen es la proyección sobre las primeras $k$ coordenadas, $p$ es un homeomorfismo sobre su imagen. Además $g(U) = M \cap G_a$ y además las matriz formada por las primeras $k$ filas de $Jp(x)$ es la matriz identidad de orden $k$. Luego en cada punto $x \in U$ el rango de $Jp(x)$ es $k$. Por tanto, $M$ es una variedad diferenciable de clase $C^q$.

# Variedad diferenciable de $\mathbb{R}^N$

2) $\Rightarrow$ 3) Por hipótesis, para cada punto $a \in M$ existe un abierto $U \subset \mathbb{R}^k$ y un abierto $G_a \subset \mathbb{R}^N$ tal que $a \in G_a$ y una aplicación $\varphi : U \longrightarrow \mathbb{R}^{N-k}$ de clase $C^q$ tal que

$$\{(x, \varphi(x)) : x \in U\} = G_a \cap M, \tag{1}$$

salvo una permutación de las coordenadas.
Definimos $F : G_a \cap (U \times \mathbb{R}^{N-k}) \longrightarrow \mathbb{R}^{N-k}$ por

$$F(z) = \varphi(Q(z)) - P(z) \quad (z \in G_a),$$

donde $Q$ es la proyección en $\mathbb{R}^N$ sobre las primeras $k$ coordenadas y $P$ es la proyección sobre las últimas $N-k$ coordenadas. Como $P$ y $Q$ son lineales, son de clase $C^\infty$, luego $F$ es de clase $C^q$. Nótese que el dominio de $F$ es un abierto que contiene al punto $a$. Si $z \in G_a \cap (U \times \mathbb{R}^{N-k})$ y $F(z) = 0$, entonces si $z = (x, y) \in U \times \mathbb{R}^{N-k}$ tenemos que

$$0 = F(z) = \varphi(x) - y,$$

luego $z = (x, \varphi(x)) \in G_a \cap M$. La igualdad (1) permite comprobar que $G_a \cap M = \{(x, y) \in G_a \cap (U \times \mathbb{R}^{N-k}) : F(x, y) = 0\}$.

# Variedad diferenciable de $\mathbb{R}^N$

Además

$$JF(x,y) = \Big( C(x,y) | -I_{N-k} \Big), \qquad \forall (x,y) \in G_a \cap (U \times \mathbb{R}^{N-k}),$$

donde $C(x,y)$ es una matriz de orden $(N-k) \times k$ y $I_{N-k}$ es la matriz identidad de orden $N-k$, ya que $F$ es suma de un campo escalr que depende sólo de las $k$ primeras coordenaadas y la opuesta de la proyección sobre las últimas $N-k$ coordenadas, luego el rango de $JF(x,y)$ es $N-k$.

3) $\Rightarrow$ 2) Se obtiene aplicando el Teorema de la función implícita. $\square$

# Variedad diferenciable de $\mathbb{R}^N$

## Ejemplos

**1)** Consideramos la siguiente ecuación

$$f(x,y) := ax + by - c = 0,$$

donde $a, b$ y $c$ son reales fijos tales que $a^2 + b^2 \neq 0$. Como $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ es una función de clase $C^\infty$ y

$$\nabla f(x,y) = (a,b).$$

luego el rango del gradiente es 1 en cada punto de $\mathbb{R}^2$, el conjunto de ceros de $f$ (una recta) es una variedad diferenciable de clase $C^\infty$ y dimensión 1 (=número de variables - número de ecuaciones).

# Variedad diferenciable de $\mathbb{R}^N$

**2)** Consideramos $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, que es el conjunto de ceros de

$$f(x, y) = x^2 + y^2 - 1.$$

Si consideramos $G = \mathbb{R}^2 \backslash \{0\} \longrightarrow \mathbb{R}$, es claro que $f \in C^\infty(G)$ y

$$\nabla f(x, y) = (2x, 2y), \quad \forall (x, y) \in \mathbb{R}^2 \backslash \{0\}.$$

Además en puntos de $G$ el gradiente de $f$ tiene rango 1. Luego $S^1$ es una variedad diferenciable de clase $C^\infty$ y dimensión 1.

# Espacios tangente y normal a una variedad

## Definición

Sean $k, N \in \mathbb{N}$ con $k < N$, $M$ una variedad de $\mathbb{R}^N$ de dimensión $k$ y $a$ un punto de $M$.
Se dice que $u \in \mathbb{R}^N$ es un **vector tangente a $M$ en $a$** si existe una curva contenida en $M$ que pasa por el punto $a$ y cuya derivada en el valor del parámetro correspondiente al punto $a$ es el vector $u$, es decir,

$$\exists \delta > 0, \exists \gamma : ]-\delta, \delta[ \longrightarrow \mathbb{R}^N \text{ continua tal que } \gamma(]-\delta, \delta[ \subset M \text{ con } \begin{cases} \gamma(0) = a \\ \gamma'(0) = u \end{cases}.$$

Se define el **espacio tangente $T_M(a)$ a la variedad $M$ en $a$** como el conjunto de vectores tangentes a $M$ en el punto $a$.
Más adelante comprobaremos que $T_M(a)$ es un subespacio vectorial.
El complemento ortogonal en $\mathbb{R}^N$ del espacio tangente en $a$, que notaremos $T_M(a)^\perp$, se llama **espacio normal a $M$ en $a$**, esto es,

$$T_M(a)^\perp = \{x \in \mathbb{R}^N : < x|u >= 0, \ \forall u \in T_M(a)\}.$$

# Espacios tangente y normal a una variedad

A la variedad afín $a + T_M(a)$ se le llama la **variedad afín tangente** a la variedad $M$ en el punto $a$. El conjunto $a + T_M(a)^{\perp}$ se le llama la **variedad afín normal** a la variedad $M$ en el punto $a$.

El siguiente resultado describe estos espacios en términos de la función cuyos ceros determina localmente la variedad en un entorno de $a$ y de una parametrización local de la variedad en $a$.

# Espacios tangente y normal a una variedad

## Proposición

Sean $k, m, N \in \mathbb{N}$ con $m + k = N$, $M$ una variedad en $\mathbb{R}^N$ de dimensión $k$, $a$ un punto de $M$ y $T_M(a)$ el espacio tangente a $M$ en $a$.

**1)** Si $F$ es una función que determina (localmente) la variedad $M$ en el punto $a$, entonces

$$T_M(a) = \operatorname{Ker} DF(a),$$

con lo que $T_M(a)$ es un subespacio vectorial de dimensión $k$, y

$$a + T_M(a) = \{x \in \mathbb{R}^N : \; DF(a)(x - a) = 0\}.$$

Además

$$\{\nabla F_1(a), \ldots, \nabla F_m(a)\}$$

es una base del espacio vectorial normal $T_M(a)^\perp$.

**2)** Si $(\Omega, U, p)$ es una parametrización cartesiana local de $M$ en $a$ y $u_0$ el único punto de $U$ tal que $p(u_0) = a$, entonces

$$T_M(a) = Dp(u_0)(\mathbb{R}^k).$$

# Espacios tangente y normal a una variedad

Los vectores columna de la matriz $Jp(u_0)$ son una base de $T_M(a)$ y por tanto

$$a + T_M(a)^{\perp} = \{x \in \mathbb{R}^N : \ \langle D_j p(u_0) | x - a \rangle = 0, \ \forall j = 1, \ldots, k\}.$$

**Demostración:** Probaremos en primer lugar que

$$Dp(u_0)(\mathbb{R}^k) \subseteq T_M(a) \subseteq \operatorname{Ker} DF(a). \tag{2}$$

Empezamos probando la inclusión $Dp(u_0)(\mathbb{R}^k) \subseteq T_M(a)$. Sea $y$ un vector de $Dp(u_0)(\mathbb{R}^k)$. Tomemos $x \in R^k$ tal que $Dp(u_0)(x) = y$. Puesto que $U$ es abierto y $u_0 \in U$, existe $\delta > 0$ tal que $]u_0 - \delta x, u_0 + \delta x[ \subset U$. Consideremos la función $\gamma : ]-\delta, \delta[ \longrightarrow \mathbb{R}^N$ definida por

$$\gamma(t) = p(u_0 + tx).$$

Puesto que $p$ es diferenciable y con imagen contenida en $M$, se sigue que $\gamma$ es una curva contenida en $M$ y con derivada en todo punto. Además

$$\begin{cases} \gamma(0) = p(u_0) = a \\ \gamma'(0) = D\gamma(0)(1) = Dp(u_0)(x) = y. \end{cases}$$

Hemos probado la primera inclusión.

# Espacios tangente y normal a una variedad

Probamos ahora la inclusión $T_M(a) \subseteq \operatorname{Ker} DF(a)$. Sea $u \in T_M(a)$ y tomemos $\delta > 0$ y $\gamma : ]-\delta, \delta[ \longrightarrow M$ continua con $\begin{cases} \gamma(0) = a \\ \gamma'(0) = u \end{cases}$. Puesto que $\gamma\big(]-\delta,\delta[\big) \subset M$, se tiene que la función $F \circ \gamma$ está definida y es nula en un entorno de 0, luego el carácter local de la diferenciabilidad nos asegura que

$$0 = D(F \circ \gamma)(0) = DF(a) \circ D\gamma(0),$$

y por tanto

$$0 = (DF(a) \circ D\gamma(0))(1) = DF(a)\big(D\gamma(0)(1)\big) = DF(a)\big(\gamma'(0)\big) = DF(a)(u).$$

Acabamos de probar la segunda inclusión.
Ahora, para que se de la igualdad en (2) nos basta observar que los espacios (vectoriales) de los extremos tienen la misma dimensión. En efecto,

$$\dim\big(\operatorname{Ker} DF(a)\big) = N - \dim\big(DF(a)(\mathbb{R}^N)\big) = N - m = k =$$

$$\operatorname{rango}\big(Jp(u_0)\big) = \dim\big(Dp(u_0)(\mathbb{R}^k)\big).$$

Hemos probado que

$$Dp(u_0)(\mathbb{R}^k) = T_M(a) = \operatorname{Ker} DF(a).$$

# Espacios tangente y normal a una variedad

Sabemos que $T_M(a) = \operatorname{Ker} DF(a)$, donde $F$ es la función cuyos ceros determina localmente la variedad en un entorno de $a$. Por tanto,

$$0 = DF(a)(x) = JF(a)x^t, \qquad \forall x \in T_M(a).$$

Luego $< \nabla F_i(a), x >= 0$ para cada $1 \leq i \leq m = N - k$. Por tanto, los vectores $\{\nabla F_i(a) : 1 \leq i \leq m\} \subset T_M(a)^\perp$. Como $\dim T_M(a) = k$, entonces $\dim T_M(a)^\perp = N - k = m$. Además, como el rango de $JF(a)$ es $m$, las $m$ filas de $JF(a)$ son linealmente independientes, es decir, los vectores de $\{\nabla F_i(a) : 1 \leq i \leq m\}$ son linealmente independientes, luego son una base de $T_M(a)^\perp$.
La igualdad

$$a + T_M(a) = \{x \in \mathbb{R}^N : DF(a)(x - a) = 0\} \tag{3}$$

es consecuencia inmediata de la igualdad $T_M(a) = \operatorname{Ker} DF(a)$, ya que si $x = a + z$, son $z \in T_M(a)$, entonces $x - a = z \in T_M(a)$, luego $0 = DF(a)(z) = DF(a)(x - a)$. En cada paso del argumento anterior, es cierto el recíproco, con lo que se obtiene la descripción (3) de $a + T_M(a)$.

# Espacios tangente y normal a una variedad

Como hemos probado que

$$T_M(a) = Dp(u_0)(\mathbb{R}^k),$$

donde $p$ es la parametrización local de $M$ en $a$ y $p(u_0) = a$ y $\dim T_M(a) = k$, entonces los vectores

$$\{Dp(u_0)(e_i) : 1 \leq i \leq k\}$$

son una base de $T_M(a)$. Como $Dp(u_0)(e_i) = Jp(u_0)(e_i)$ es la columna $i-$ésima de la matriz $Jp(u_0)$, entonces los vectores columna de la matriz $Jp(u_0)$ forman una base de $T_M(a)$. Notamos por $\{D_j p(u_0) : 1 \leq j \leq k\}$ a los vectores columna anteriores, luego

$$T_M(a)^{\perp} = \{x \in \mathbb{R}^N :< D_j p(u_0), x >= 0, \forall 1 \leq j \leq k\}.$$

Como consecuencia,

$$a + T_M(a)^{\perp} = \{x \in \mathbb{R}^N :< D_j p(u_0), x - a >= 0, \forall 1 \leq j \leq k\}.$$

# Espacios tangente y normal a una variedad

## Ejemplos

**1)** Consideramos la siguiente ecuación

$$f(x,y) := bx + cy - d = 0,$$

donde $b, c$ y $d$ son reales fijos tales que $b^2 + c^2 \neq 0$. Sabemos que es una variedad diferenciable de dimensión 1. En este caso $\nabla f(x,y) = (b,c)$ para cada punto $(x,y) \in \mathbb{R}^2$. Por tanto, si $a = (x_0, y_0)$ verifica que $f(a) = 0$, tenemos que

$$T_M(a) = \{(x,y) \in \mathbb{R}^2 :< \nabla f(a), (x,y) = 0 >\} =$$

$$\{(x,y) \in \mathbb{R}^2 : bx + cy = 0\}.$$

Por tanto $T_M(a)^\perp$ está generado por $(b,c)$. Luego

$$a + T_M(a) = \{(x,y) \in \mathbb{R}^2 :< (b,c), (x - x_0, y - y_0) >= 0\}.$$

# Espacios tangente y normal a una variedad

Como

$$0 =< (b, c), (x - x_0, y - y_0) >= b(x - x_0) + c(y - y_0) =$$

$$bx + cy - (bx_0 + cy_0) = bx + cy - d.$$

Entonces, en este caso

$$a + T_M(a) = M, \quad \forall a \in M,$$

mientras que $T_M(a)$ es el subespacio vectorial asociado a $M$.

En general, si $M$ es una variedad afín, en cada punto de $M$ la variedad afín tangente en ese punto coincide con $M$ y el espacio vectorial tangente en cada punto es el subespacio vectorial asociado a $M$.

# Espacios tangente y normal a una variedad

**2)** Consideramos $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, que es el conjunto de ceros de

$$f(x, y) = x^2 + y^2 - 1.$$

En este caso $\nabla f(x, y) = (2x, 2y)$. Por tanto, si $a = (x_0, y_0) \in S^1$,

$$T_{S^1}(a) = \{(x, y) \in \mathbb{R}^2 : x_0 x + y_0 y = 0\},$$

y el vector $(x_0, y_0)$ genera $T_{S^1}(a)$. La variedad afín tangente se obtiene trasladando por $a$ el subespacio $T_{S^1}(a)$.

**3)** Por último, si $\varphi : \mathbb{R}^N \longrightarrow \mathbb{R}$ es un campo escalar $C^1$, el conjunto $M = \{(x, \varphi(x)) : x \in \mathbb{R}^N\} \subset \mathbb{R}^{N+1}$ es una variedad diferenciable de dimensión $N$ y clase $C^1$. En este caso, la aplicación $p : \mathbb{R}^N \longrightarrow \mathbb{R}^{N+1}$ dada por

$$p(x) = (x, \varphi(x)) \quad (x \in \mathbb{R}^N)$$

es una parametrización local para cada punto de la variedad. Si $a \in \mathbb{R}^N$, entonces el elemento $(a, \varphi(a) \in M.$ )

$$T_M((a, \varphi(a)) = Dp(a)(\mathbb{R}^N).$$

# Espacios tangente y normal a una variedad

Luego

$$(a,\varphi(a)) + T_M((a,\varphi(a)) = \{y \in \mathbb{R}^{N+1} : y - (a,\varphi(a)) \in Dp(a)(\mathbb{R}^N)\}.$$

Si $x \in \mathbb{R}^N$ tenemos

$$Dp(a)(x) = Jp(a)x^t = (x_1, x_2, \cdots x_N, < \nabla\varphi(a), x >)$$

Luego $y$ pertenece a la variedad afín tangente a $M$ en $(a,\varphi(a))$ sí, y sólo si,

$$y - (a,\varphi(a)) = (x_1, x_2, \cdots x_N, < \nabla\varphi(a), x >),$$

para algún vector $x \in \mathbb{R}^N$. Luego

$$y_j - a_j = x_j, \forall j \le N \quad \text{y} \quad y_{N+1} - \varphi(a) =< \nabla\varphi(a), (y_1 - a_1, y_2 - a_2, \cdots, y_N - a_N) >$$

En definitiva, la variedad afín tangente a $M$ en $(a,\varphi(a))$ viene dada por

$$\{y \in \mathbb{R}^{N+1} : y_{N+1} = \varphi(a) + \sum_{j=1}^{N} D_j\varphi(a)(y_j - a_j)\},$$

que es, por definición, el hiperplano tangente a la gráfica de $\varphi$ en el punto $a$.

# Extremos condicionados

## Definición

Sean $\Omega \subset \mathbb{R}^N$, $f : \Omega \longrightarrow \mathbb{R}$ y $M \subset \Omega$.
Se dice que $f$ tiene en un punto $a \in M$ un **máximo local condicionado por** $M$ si existe $r > 0$ tal que

$$f(x) \leq f(a), \qquad \forall x \in M \cap B(a, r).$$

Si para algún positivo $r > 0$ se verifica tal que

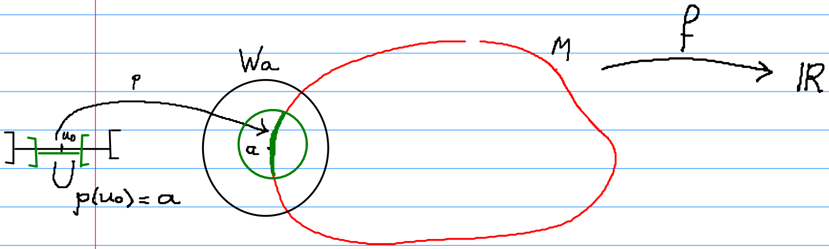
$$f(x) \geq f(a), \qquad \forall x \in M \cap B(a, r).$$

y $a \in M$, diremos que $f$ tiene en $a$ un **mínimo local condicionado por** $M$.
En cualquiera de los dos casos, se dice que $f$ tiene en $a$ un **extremo local condicionado por** $M$**.**

# Extremos condicionados en una variedad

**Idea intuitiva**

$\boxed{k=1,\ N=2}$



# Variedades diferenciables

En lo que sigue cada vez que mencionemos que "$M$ es una variedad diferenciable" entenderemos que $M \subset \mathbb{R}^N$ es una variedad diferenciable de clase $C^1$ y que $k$ es la dimensión de la variedad, donde $1 \leq k < N$. Además, dado un punto $a \in M$, notaremos por $F$ al campo vectorial definido en un entorno abierto de $a$ cuyos ceros determinan localmente la variedad, mientras que $p$ será la parametrización local de $M$ en $a$.
Si $p(u_0) = a$, sabemos que $\dim T_M(a) = k$ y además

$$T_M(a) = \operatorname{Ker} DF(a) = Dp(u_0)(\mathbb{R}^k).$$

Por último, los vectores fila de $DF(a)$ son una base de $T_M^{\perp}$ y los vectores columna de $Jp(u_0)$ son una base de $T_M(a)$.

# Extremos condicionados

A continuación trasladaremos el estudio de extremos condicionados por una variedad al de extremos relativos de un campo escalar definido en un abierto de $\mathbb{R}^k$.

## Proposición

Sea $M$ una variedad diferenciable, $M \subset \Omega \subset \mathbb{R}^N$, tal que $\Omega$ es abierto, y $f : \Omega \longrightarrow \mathbb{R}$ un campo escalar. Equivalen

1) $f$ tiene en un punto $a \in M$ un extremo local condicionado por $M$.
2) $f \circ p$ alcanza en $u_0$ un extremo relativo en $u_0$, donde $p$ es la parametrización local de $M$ en $a$ y $p(u_0) = a$.

Además, en caso de que se verifiquen las afirmaciones anteriores el extremo condicionado de $f$ es del mismo tipo que el extremo relativo de $f \circ p$, es decir ambas funciones tienen un mínimo o ambas tienen un máximo.

Como habitualmente las variedades vendrán dadas como conjunto de ceros de un campo vectorial $(F)$, daremos una formulación que sea útil para calcular los extremos de un campo escalar condicionados por una variedad diferenciable.

# Función de Lagrange

## Problema

Supongamos que $M$ es una variedad diferenciable, $f : \Omega \longrightarrow \mathbb{R}$ y $M \subset \Omega$, donde $\Omega$ es un abierto de $\mathbb{R}^N$. Queremos estudiar los extremos locales de $f$ condicionados por $M$.

Llamaremos **función de Lagrange** (depende de $M$ y de $f$) al campo escalar definido por

$$L(x, \lambda) = f(x) + \sum_{i=1}^{N-k} \lambda_i F_i(x) \qquad (x \in \Omega, \lambda \in \mathbb{R}^{N-k}).$$

A las componentes del vector $\lambda$ se llaman **multiplicadores de Lagrange.**

El siguiente resultado da una condición necesaria para que un punto de la variedad sea un extremo de $f$ condicionado por $M$, en caso de que $f$ sea diferenciable.

### Teorema de Lagrange

Sean $M$ una variedad diferenciable en $\mathbb{R}^N$ de dimensión $k$, $\Omega$ un abierto de $\mathbb{R}^N$ que contiene a $M$ y $f:\Omega\longrightarrow\mathbb{R}$ un campo escalar diferenciable. Si la función $f$ alcanza en un punto $a\in M$ un extremo condicionado por $M$, existe un único $\lambda\in\mathbb{R}^{N-k}$ tal que

$$\nabla L(a,\lambda)=0,$$

donde $L$ es la función de Lagrange asociada al problema anterior y $F$ es el campo vectorial cuyos ceros determinan localmente a $M$ en un entorno de $a$. Esto es,

$$L(x,\lambda)=f(x)+\sum_{i=1}^{N-k}\lambda_i F_i(x)\qquad (x\in\Omega,\lambda\in\mathbb{R}^{N-k}).$$

# Extremos condicionados

**Demostración:** Por hipótesis $f : \Omega \longrightarrow \mathbb{R}$ alcanza un extremo local en $a \in M$ condicionado por $M$, luego $f \circ p$ alcanza en $u_0$ un extremo relativo, donde $p : U \subset \mathbb{R}^k \longrightarrow \mathbb{R}^N$ es la parametrización local de $f$ en $a$ y $p(u_0) = a$. Como $p$ y $f$ son diferenciables, entonces $f \circ p$ también lo es, luego $0 = D(f \circ p)(u_0) = Df(a) \circ Dp(u_0)$. Esto es, $\nabla f(a) Jp(u_0) = 0$. Por tanto, $< \nabla f(a), Djp(u_0) >= 0$ para cada $1 \leq j \leq k$. Como los vectores

$$\{D_j p(u_0) : 1 \leq j \leq k\}$$

generan el espacio $T_M(a)$, entonces $\nabla f(a) \in T_M(a)^{\perp}$.
Sabemos que

$$\{\nabla F_j(a) : 1 \leq j \leq N - k\}$$

es una base del espacio $T_M(a)^{\perp}$, por tanto, existen un único $\alpha \in \mathbb{R}^{N-k}$, tal que

$$\nabla f(a) = \sum_{j=1}^{N-k} \alpha_j \nabla F_j(a).$$

# Extremos condicionados

Si tomamos $\lambda = -\alpha \in \mathbb{R}^{N-k}$, entonces se verifica que

$$\nabla f(a) + \sum_{j=1}^{N-k} \lambda_j \nabla F_j(a) = 0 \tag{4}$$

y como $a \in M$, dado que $F$ determina localmente la variedad en un entorno de $a$ se verifica que

$$F_j(a) = 0, \qquad \forall 1 \leq j \leq N-k. \tag{5}$$

Por las igualdades (4) y (5) sabemos que el vector $(a, \lambda)$ es solución del sistema de Lagrange

$$\nabla L(x, \lambda) = 0,$$

donde

$$L(x, \lambda) = f(x) + \sum_{j=1}^{N-k} \lambda_j F_j(x).$$

Además $\lambda$ es único por serlo $\alpha$. $\square$

# Extremos condicionados

## Ejemplo

Calcula el máximo y el mínimo absolutos de la función $f(x, y) = xy$ en $S^1$, donde

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

Como $f$ es continua y $S^1$ es un compacto, por el Teorema de Weiorstrass, $f$ alcanza el mínimo y el máximo y absolutos en $S^1$. Necesariamente éstos han de ser extremos de $f$ condicionados por $S^1$. Es claro que $f$ es diferenciable. En este caso, la función $F$ cuyos ceros determinan la variedad viene dada por

$$F(x, y) = x^2 + y^2 - 1,$$

definida en $W = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \neq 0\}$. Sabemos que $S^1$ es una variedad diferenciable.
La función de Lagrange asociada al problema viene dada por

$$L(x, y, \lambda) = f(x, y) + \lambda(x^2 + y^2 - 1) = xy + \lambda(x^2 + y^2 - 1).$$

Es claro que

$$\nabla L(x, y, \lambda) = (y + 2\lambda x, x + 2\lambda y, x^2 + y^2 - 1).$$

# Extremos condicionados

Los puntos críticos de $L$ son soluciones del sistema de ecuaciones

$$y + 2\lambda x = 0, \quad x + 2\lambda y = 0, \quad x^2 + y^2 = 1.$$

Luego $y = -2\lambda x$. Sustituyendo en la segunda ecuación tenemos $x + 2\lambda(-2\lambda x) = 0$, es decir $x(1 - 4\lambda^2) = 0$. Si $x = 0$, entonces $y = \pm 1$ y $\lambda = 0$. Pero entonces no se verifica la primera ecuación.
En otro caso, $1 = 4\lambda^2$, esto es, $\lambda^2 = \frac{1}{4}$, tenemos $\lambda = \pm\frac{1}{2}$. Si $\lambda = \frac{1}{2}$, sustituyendo en las ecuaciones obtenemos

$$x + y = 0, \quad x^2 + y^2 = 1,$$

luego las soluciones son

$$\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{2}\right), \quad \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{2}\right).$$

Si $\lambda = -\frac{1}{2}$, haciendo el mismo procedimiento se obtienen las soluciones

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{2}\right), \quad \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{1}{2}\right).$$

# Extremos condicionados

Por último evaluamos en los puntos de la variedad que han parecido al resolver el sistema de Lagrange

$$f\Big(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\Big)=f\Big(-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\Big)=\frac{1}{2}$$

y

$$f\Big(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\Big)=f\Big(-\frac{1}{\sqrt{2}},+\frac{1}{\sqrt{2}}\Big)=-\frac{1}{2}.$$

Luego el máximo absoluto de $f$ sobre $S^1$ vale $\frac{1}{2}$ y el mínimo absoluto $-\frac{1}{2}$. Cada uno de los dos valores se alcanza en dos puntos de la esfera.

# Extremos condicionados

Ahora recurriremos al criterio que existe en términos del hessiano para clasificar los extremos relativos para obtener un criterio análogo para los extremos condicionados por una variedad.

## Teorema

Sean $M$ una variedad diferenciable en $\mathbb{R}^N$ de dimensión $k$, y clase $C^2$, $\Omega$ un abierto de $\mathbb{R}^N$ que contiene a $M$ y $f : \Omega \longrightarrow \mathbb{R}$ un campo escalar de clase $C^2$. Supongamos que $(a, \lambda)$ es solución del sistema de Lagrange asociado a $M$ y a $f$. Definimos

$$L_\lambda(x) = L(x, \lambda)$$

y llamamos $Q$ a la restricción del hessiano de $L_\lambda$ en $a$ al espacio $T_M(a)$.
Se verifica que
**1)** Si $Q$ es definida negativa, entonces $f$ tiene en $a$ un máximo local condicionado por $M$.
**2)** Si $Q$ es definida positiva, entonces $f$ tiene en $a$ un mínimo local condicionado por $M$.

# Extremos condicionados

**3)** Si $f$ tiene en $a$ un máximo local condicionado por $M$, $Q$ es semidefinida negativa.
**4)** Si $f$ tiene en $a$ un mínimo local condicionado por $M$, $Q$ es semidefinida positiva.
**5)** Como consecuencia, si $Q$ es indefinida, entonces $f$ no tiene un extremo local en $a$ condicionado por $M$.

En este caso no se incluye la demostración. Notamos en primer lugar que si $(a, \lambda)$ es un punto crítico de la función de Lagrange, entonces el elemento $u_0$ que se aplica en $a$ mediante la parametrización local es un punto crítico de $f \circ p$. En la prueba se usa la equivalencia entre los extremos relativos de $f \circ p$ y los extremos locales condicionados por $M$, y el hecho de que el hessiano de $f \circ p$ en $u_0$ coincide con $Q$. Por último, el criterio se obtiene a partir de las condiciones necesarias y suficientes para extremos relativos.

# Extremos condicionados

## Ejemplo

Determina las dimensiones del ortoedro en $\mathbb{R}^3$ de volumen fijo una unidad y superficie lateral mínima.

Llamamos $x, y, z$ a las medidas de los lados del ortoedro. Imponemos que el volumen sea 1 y optimizamos la (mitad de la) superficie de los lados. Consideramos

$$M = \{(x, y, z) \in G : xyz = 1\} \subset \mathbb{R}^3,$$

donde $G = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$, luego

$$F(x, y, z) = xyz - 1.$$

es la función cuyos ceros define a $M$, que es de clase $C^\infty$ y verifica

$$\nabla F(x, y, z) = (yz, xz, xy),$$

luego el rango de $\nabla F$ es 1 en cada punto de $G$. Por tanto, $M$ es una variedad diferenciable de dimensión 2.
Queremos minimizar en $M$ la función $f : G \longrightarrow \mathbb{R}$ definida por

$$f(x, y, z) = xy + xz + yz.$$

# Extremos condicionados

La función de Lagrange $L$ en esta caso está definida en $L : G \times \mathbb{R} \longrightarrow \mathbb{R}$ por

$$L(x, y, z, \lambda) = xy + xz + yz + \lambda(xyz - 1).$$

Calculamos el gradiente de $L$, ya que hemos de resolver el sistema de Lagrange, dado por

$$\left\{ \begin{array}{c} y + z + \lambda yz = 0 \\ x + z + \lambda xz = 0 \\ x + y + \lambda xy = 0 \\ xyz = 1 \\ x, y, z > 0 \end{array} \right.$$

Multiplicando la primera ecuación por $x$, la segunda por $y$ y la tercera por $z$ obtenemos

$$\left\{ \begin{array}{c} xy + xz + \lambda xyz = 0 \\ xy + yz + \lambda xyz = 0 \\ xz + yz + \lambda xyz = 0 \end{array} \right.$$

Por tanto $xy + xz = xy + yz = xz + yz$, luego

$$(x - y)z = 0 = x(y - z) \Rightarrow x = y = z,$$

donde hemos usado que $x, z \neq 0$.

# Extremos condicionados

Como el volumen es 1, la única solución es la dada por

$$(x, y, z, \lambda) = (1, 1, 1, -2).$$

Vamos a usar el criterio que usa el hessiano, aunque la información que se obtiene es local.
Fijamos $\lambda = -2$ y consideramos la función

$$L_{-2}(x, y, z) = L(x, y, z, -2) = xy + yz + xz - 2(xyz - 1).$$

Luego

$$\nabla L_{-2}(x, y, z) = (y + z - 2yz, x + z - 2xz, x + y - 2xy).$$

Como consecuencia,

$$HL_{-2}(x, y, z) = \begin{pmatrix} 0 & 1 - 2z & 1 - 2y \\ 1 - 2z & 0 & 1 - 2x \\ 1 - 2y & 1 - 2x & 0 \end{pmatrix}$$

# Extremos condicionados

Evaluando en el único punto crítico tenemos

$$H := HL_{-2}(1,1,1) = \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$

Ahora calculamos $T_M(1,1,1)$ y restringimos la forma cuadrática asociada a la matriz anterior al subespacio tangente. Sabemos que

$$T_M(1,1,1) = \mathrm{Ker}\, DF(1,1,1) = \{(x,y,z) \in \mathbb{R}^3 : x+y+z=0\},$$

donde hemos usado que $\nabla F(1,1,1) = (1,1,1)$. Por tanto, los vectores $\{(1,-1,0),(0,1,-1)\}$ son una base de $T_M(1,1,1)$. Obtenemos ahora la matriz de la forma cuadrática asociada a $H$ restringida a $TM(1,1,1)$ en función de la base anterior. Para ello calculamos $K^t HK$, donde

$$K = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

# Extremos condicionados

En este caso

$$K^t H K = K^t \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Luego la forma cuadrática asociada a la restricción del hessiano de $L_{-2}(1,1,1)$ a $T_M(1,1,1)$ es definida positiva, luego $f$ tiene en $(1,1,1)$ un mínimo local condicionado por $M$.

Para probar que la mínima superfice lateral se alcanza en $(1,1,1)$ puede usarse la desigualdad de las medias, aplicada a los reales $xy, yz, xz$, luego

$$1 = \sqrt[3]{x^2y^2z^2} \leq \frac{xy+yz+xz}{3}.$$

Nótese que para $x = y = z = 1$ se da la igualdad, luego la superficie lateral del ortoedro es mínima para los valores $x = y = z = 1$.