CÁLCULO VECTORIAL

TEMA 2

EXTREMOS ABSOLUTOS Y RELATIVOS

Recordemos un importante resultado que vimos en el curso de CDIVV, conocido como el *Teorema de Weierstrass*. Dado un campo escalar continuo $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ sobre un conjunto compacto U (es decir, U es cerrado y acotado), entonces existen puntos $\boldsymbol{a}, \boldsymbol{b} \in U$ tales que $f(\boldsymbol{b}) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{a})$, para todo $\boldsymbol{x} \in U$. Al valor $f(\boldsymbol{a})$ se le llama *máximo absoluto o global de* f, alcanzado en el punto $\boldsymbol{a}$ (aunque podría alcanzarse en más de un punto del dominio de f). De manera análoga, al valor $f(\boldsymbol{b})$ se le llama *mínimo absoluto o global de* f.

Tal como ocurría con funciones de una variable, hallar los extremos absolutos y relativos de un campo escalar nos ayudará a conocer parte de su comportamiento y la forma aproximada que tiene su gráfica. El problema es que esto no siempre resulta sencillo de hacer para cualquier campo. Sin embargo, pidiendo algunas condiciones que tienen que ver con diferenciabilidad, es posible demostrar algunos criterios que permitan encontrar los extremos de una función al menos de forma local.

2.1 Definición y ejemplos de extremos absolutos y relativos

Comencemos con las definiciones de extremos absolutos.

Definición 2.1.1. *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar con dominio U, y sean $M, m \in \mathbb{R}$.*

1. *Se dice que M es el* ***máximo absoluto*** *o* ***global*** *de f si existe $\boldsymbol{x_0} \in U$ tal que $M = f(\boldsymbol{x_0})$ y $f(\boldsymbol{x}) \leq f(\boldsymbol{x_0}) = M$, para todo $\boldsymbol{x} \in U$. En este caso, se dice que f alcanza su máximo absoluto en $\boldsymbol{x_0} \in U$.*

2. *De manera similar, se dice que m es el **mínimo absoluto** o **global** de f si existe $\boldsymbol{x_0} \in U$ tal que $m = f(\boldsymbol{x_0})$ y $m = f(\boldsymbol{x_0}) \leq f(\boldsymbol{x})$, para todo $\boldsymbol{x} \in U$. En este caso, se dice que f alcanza su mínimo absoluto en $\boldsymbol{x_0} \in U$.*

*Por **extremo absoluto** o **global** de f nos referiremos a un valor real que es máximo o mínimo absoluto de f.*

Observación 2.1.2. *Tenemos algunos comentarios sobre la definición anterior:*

1. *El máximo y el mínimo absoluto de un campo escalar f, en caso de existir, son únicos. Lo que no es necesariamente único son los puntos donde se alcanzan. Por ejemplo, la función constante $f(x, y) = 1$ tiene máximo y mínimo absolutos, a saber, el valor 1, y los alcanza en cualquier punto de su dominio. Por otro lado, $g(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ posee mínimo absoluto igual a 0, y lo alcanza únicamente en $(0, 0, 0)$. Sin embargo, dicha función no posee máximo absoluto por no ser acotada en su dominio (= $\mathbb{R}^3$).*

2. *Los máximos y mínimos absolutos de una función dependen del dominio donde se les considere. Si tomamos la función $g(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ del punto anterior, sabemos que no posee máximo absoluto si se considera su dominio maximal, es decir, $\mathbb{R}^3$. Sin embargo, si consideramos la restricción $g|_{\overline{B}((0,0,0),1)} \colon \overline{B}((0, 0, 0), 1) \to \mathbb{R}$, donde $\overline{B}((0, 0, 0), 1)$ denota la bola cerrada de centro $(0, 0, 0)$ y radio 1, entonces 1 es el máximo absoluto de $g|_{\overline{B}((0,0,0),1)}$, y lo alcanza en cualquier punto de la frontera de dicha bola, es decir, en la esfera unitaria de centro $(0, 0, 0)$, que se denota por S^2.*

3. *La nociones de extremos absolutos y relativos (que presentaremos más adelante) no tienen sentido para campos vectoriales $f \colon U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, ya que en $\mathbb{R}^m$ con $m \geq 2$ no tenemos una relación de orden (o al menos no una que se comporte tan bien como el orden definido para los números reales).*

4. *Tenga en cuenta que los conceptos de extremos absolutos y relativos de una función no dependen de que ésta sea continua o diferenciable. Sin embargo, estos dos atributos pueden garantizar ciertos hechos. Ya mencionamos que la continuidad sobre un conjunto compacto garantiza la existencia del máximo y del mínimo absoluto de una función, por el Teorema de Weierstrass. Una limitación de este teorema es que no nos dice cómo hallar estos extremos en la práctica. Por otro lado, veremos más adelante que ser diferenciable hasta orden 2 puede ayudar a encontrar extremos relativos.*

Ejemplo 2.1.3. *La función $f \colon \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = 2 - x^2 - y^2$ alcanza su máximo absoluto (= 2) en el punto $(0, 0)$. Esto lo podemos notar algebraicamente, pues $x^2 + y^2 \geq 0$ implica que $f(x, y) = 2 - x^2 - y^2 \leq 2 = f(0, 0)$, para todo $(x, y) \in \mathbb{R}^2$.*

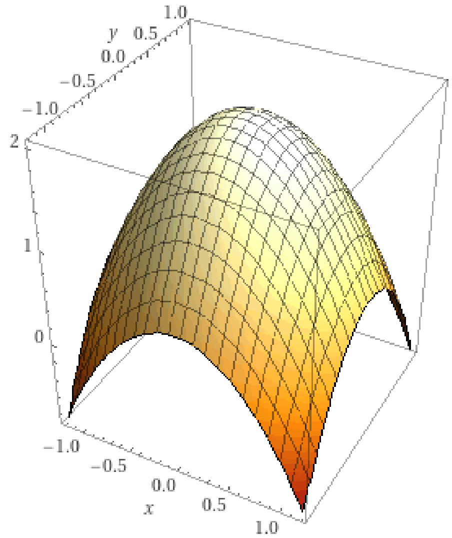

Figura 2.1: Gráfica de f que corresponde a un paraboloide circular. `(Wolfram Alpha)`

Definición 2.1.4. *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar, y sean $M, m \in \mathbb{R}$.*

1. *Se dice que M es un **mínimo relativo** o **local** de f si existen $\boldsymbol{x_0} \in U$ y $\delta > 0$ tales que $M = f(\boldsymbol{x_0})$ y $f(\boldsymbol{x}) \leq f(\boldsymbol{x_0}) = M$, para todo $\boldsymbol{x} \in U \cap B(\boldsymbol{x_0}, \delta)$. En este caso, se dice que f alcanza un máximo relativo en $\boldsymbol{x_0} \in U$.*

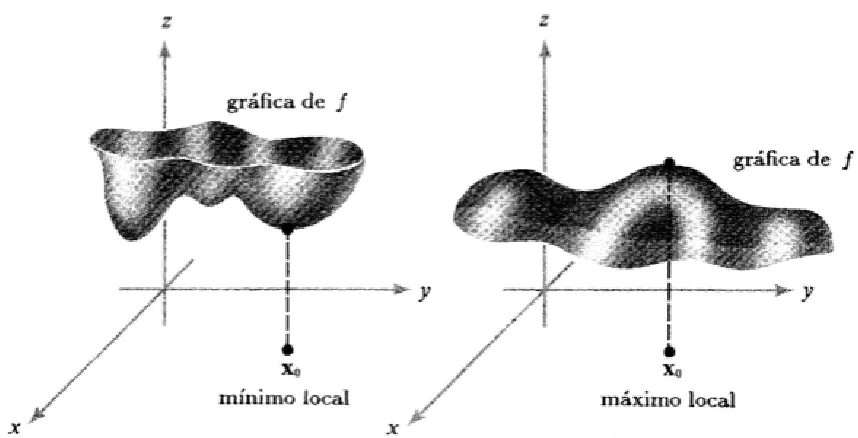

Figura 2.2: Para campos escalares de dos variables, hablando de forma intuitiva, en los puntos donde se alcanza un máximo o mínimo local se puede determinar un entorno en donde la gráfica de la función es una montaña o una depresión. Ahora, tal relieve no tiene por qué ser el más alto o profundo en la gráfica de la función. `(Imagen tomada de Cálculo Vectorial - Marsden & Tromba)`

2. *De manera similar, se dice que m es un **mínimo relativo** o **local** de f si existen $\boldsymbol{x_0} \in U$ y $\delta > 0$ tales que $m = f(\boldsymbol{x_0})$ y $m = f(\boldsymbol{x_0}) \leq f(\boldsymbol{x})$, para todo $\boldsymbol{x} \in U \cap B(\boldsymbol{x_0}, \delta)$. En este caso, se dice que f alcanza un mínimo relativo en $\boldsymbol{x_0} \in U$.*

*Por **extremo relativo** de f nos referiremos a un valor real que es máximo o mínimo relativo de f.*

Observación 2.1.5. *Tenemos algunos comentarios sobre la definición anterior:*

1. *Todo extremo absoluto de una función es también relativo. Basta con tomar cualquier $\delta > 0$ en la definición anterior. Bajando esto a tierra, literalmente, pensemos en el volcán Chimborazo (Ecuador). En su cima se registra el punto más alto de la superficie terrestre, es decir el máximo absoluto (midiendo desde el centro de la Tierra). Esta altura también es un máximo relativo si consideramos el entorno formado por el territorio de Ecuador. Por otro lado, un extremo relativo no necesariamente es absoluto, como veremos más adelante en los ejemplos. Volviendo a las analogías geográficas, la altura del monte Nanga Parbat es un máximo relativo para la superficie de la tierra, en el entorno formado por el territorio de Pakistán.*

2. *Así como pasa con los extremos absolutos, una función puede no tener ni máximos ni mínimos relativos. Por ejemplo, la función $f(x,y) = -x-y$ representa al plano $x+y+z=0$, que pasa por el origen y que no es paralelo a ninguno de los planos coordenados. Al tratarse entonces de un plano inclinado, f no posee extremos relativos (ni absolutos).*

Ejemplo 2.1.6. *La función $g\colon \mathbb{R}^2 \to \mathbb{R}$ definida por $g(x,y) = x^2 - y^2$ no alcanza máximo ni mínimo relativo en $(0,0)$. Esto se debe a que, para todo $\varepsilon > 0$, se pueden encontrar puntos $(x_1,y_1),(x_2,y_2) \in B((0,0),\varepsilon)$ tales que $g(x_1,y_1) > g(0,0) = 0 > g(x_2,y_2)$. En efecto, tomamos $\varepsilon > 0$ arbitrario y consideramos la bola abierta $B((0,0),\varepsilon)$. Dicha bola contiene a los puntos $(\varepsilon/2,0)$ y $(0,\varepsilon/2)$, para los cuales se tiene $g(\varepsilon/2,0) = \varepsilon^2/4 - 0^2 = \varepsilon^2/4 > 0$ y $g(0,\varepsilon/2) = 0^2 - \varepsilon^2/4 = -\varepsilon^2/4 < 0$.*

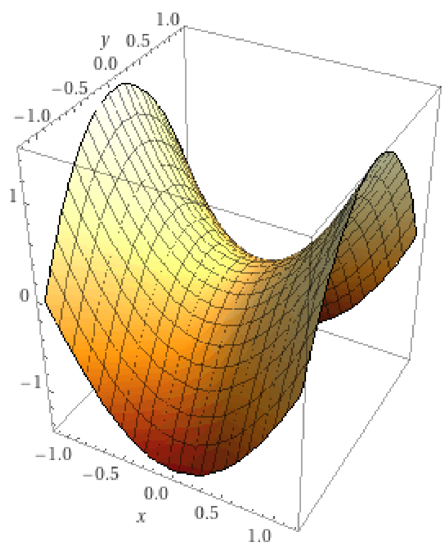

Figura 2.3: Gráfica de g que corresponde a la silla de montar. `(Wolfram Alpha)`

El objetivo principal de estas notas es entender cómo hallar y clasificar extremos relativos de un campo escalar. La teoría de funciones diferenciables, y más específicamente, los desarrollos de Taylor, nos dará las herramientas necesarias para tal fin.

Recordemos que en el curso de CDI1V se usaban los criterios de la derivada primera y segunda para clasificar extremos relativos de funciones de una variable. Para la generalización de esto que presentaremos más adelante usaremos los diferenciales de orden 1 y 2 de un campo escalar (que sea dos veces diferenciable, claro está). Nos va a ayudar un poco repasar lo que ocurre para funciones de una variable.

Proposición 2.1.7 (criterio de la derivada primera)**.** *Sea $f\colon U \subseteq \mathbb{R} \to \mathbb{R}$ una función con dominio U un intervalo, y derivable en U° (el interior de U).*

1. *Si para $c \in U^\circ$ existe $\delta > 0$ tal que $f'(x) < 0$ para $x \in U$ y $c - \delta < x < c$, y $f'(x) > 0$ para $x \in U$ y $c < x < c + \delta$, entonces f alcanza un mínimo relativo $f(c)$ en c.*

2. *Si para $c \in U^\circ$ existe $\delta > 0$ tal que $f'(x) > 0$ para $x \in U$ y $c - \delta < x < c$, y $f'(x) < 0$ para $x \in U$ y $c < x < c + \delta$, entonces f alcanza un máximo relativo $f(c)$ en c.*

Proposición 2.1.8 (criterio de la derivada segunda)**.** *Sea $f\colon U \subseteq \mathbb{R} \to \mathbb{R}$ una función con dominio U un intervalo, y derivable dos veces en U°.*

1. *Si para $c \in U^\circ$ se tiene que $f''(c) > 0$, entonces f alcanza un mínimo relativo $f(c)$ en c.*

2. *Si para $c \in U^\circ$ se tiene que $f''(c) < 0$, entonces f alcanza un máximo relativo $f(c)$ en c.*

3. *Si $f''(c) = 0$, el criterio no decide.*

2.2 Puntos críticos

En lo que sigue presentaremos los análogos de los criterios anteriores para campos escalares, en términos de los diferenciales $df_{\boldsymbol{x_0}}$ y $d^2 f_{\boldsymbol{x_0}}$ de orden 1 y 2 de $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ en puntos $\boldsymbol{x_0} \in U$ que cumplen la definición que aparece más abajo. Antes de especificar esto, es importante recordar también que una condición necesaria para que una función de una variable alcance un extremo relativo en un punto c es que su derivada se anule en c. Sabemos de CDI1V que a estos puntos se le conocen como puntos críticos. Sabemos además que en todo punto donde se alcanza un extremo relativo es un punto crítico, aunque el recíproco no siempre es cierto. Esto también pasará para campos escalares dentro del sentido dado por la siguiente definición.

Definición 2.2.1 (puntos estacionarios)**.** *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar diferenciable. Un punto $\boldsymbol{x_0} \in U$ es un* ***punto crítico*** *o* ***estacionario de $\boldsymbol{f}$*** *si $df_{\boldsymbol{x_0}} = 0$.*

*Un punto estacionario $\boldsymbol{x_0}$ de f es un **punto de silla** si para todo $\varepsilon > 0$, existen $\boldsymbol{x}, \boldsymbol{y} \in B(\boldsymbol{x_0}, \varepsilon)$ tales que $f(\boldsymbol{x}) > f(\boldsymbol{x_0}) > f(\boldsymbol{y})$. En otras palabras, un punto de silla de un campo escalar es un punto crítico donde éste no alcanza ni máximo ni mínimo relativo.*

Observación 2.2.2. *Dado un campo escalar $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable, para que un punto $\boldsymbol{x_0} \in U$ sea estacionario basta con que todas las derivadas parciales de f se anulen en $\boldsymbol{x_0}$. Efectivamente, si $\frac{\partial f}{\partial x_i}(\boldsymbol{x_0}) = 0$ para todo $1 \leq i \leq n$, entonces $\nabla f(\boldsymbol{x_0}) = \mathbf{0}$, es decir, que la matriz que representa a $df_{\boldsymbol{x_0}}\colon \mathbb{R}^n \to \mathbb{R}$ en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}$ es la matriz fila $\mathbf{0}$, por lo que la transformación lineal $df_{\boldsymbol{x_0}}$ debe ser cero.*

Ejemplo 2.2.3.

1. *El punto $(0,0)$ de la función $f(x,y) = 2-x^2-y^2$ en el Ejemplo 2.1.3 es un punto estacionario de f. En efecto, como $\frac{\partial f}{\partial x} = -2x$ y $\frac{\partial f}{\partial y} = -2y$, se tiene que $\frac{\partial f}{\partial x}(0,0) = 0$ y $\frac{\partial f}{\partial y}(0,0) = 0$. Entonces, $(0,0)$ es un punto estacionario de f.*

2. *El punto $(0,0)$ de la función $g(x,y) = x^2 - y^2$ en el Ejemplo 2.1.6 es un punto estacionario de silla de g. En este caso, $\frac{\partial g}{\partial x} = 2x$ y $\frac{\partial g}{\partial x} = -2y$, de donde $\frac{\partial g}{\partial x}(0,0) = 0$ y $\frac{\partial g}{\partial x}(0,0) = 0$.*

Vamos a concluir esta sección demostrando que, los puntos en los cuales un campo escalar diferenciable alcanza un máximo o mínimo relativo son puntos estacionarios.

Proposición 2.2.4 (los extremos relativos son puntos estacionarios)**.** *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar diferenciable y $\boldsymbol{x_0} \in U$. Si f alcanza un extremo relativo en $\boldsymbol{x_0}$, entonces $\boldsymbol{x_0}$ es un punto estacionario de f.*

Hagamos una interpretación geométrica del resultado anterior. Supongamos que tenemos un campo escalar diferenciable de dos variables $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$, y que en $\boldsymbol{x_0} \in U$ se alcanza un extremo relativo. Entonces, localmente en un entorno de $\boldsymbol{x_0}$, la gráfica de f presenta una montaña o una depresión. En dicho entorno podemos aproximar f como:

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) \approx f(\boldsymbol{x_0}) + df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) = f(\boldsymbol{x_0}) + \nabla f(\boldsymbol{x_0}) \cdot \boldsymbol{\Delta x}.$$

Es decir, que cerca del punto $(\boldsymbol{x_0}, f(\boldsymbol{x_0}))$ la gráfica de f se parece a la del plano con normal $\left(\frac{\partial f}{\partial x}(\boldsymbol{x_0}), \frac{\partial f}{\partial y}(\boldsymbol{x_0}), -1\right)$ y que pasa por $(\boldsymbol{x_0}, f(\boldsymbol{x_0}))$. Al ser $f(\boldsymbol{x_0})$ un extremo relativo alcanzado en $\boldsymbol{x_0}$, este vector normal debe ser paralelo al eje Z, por lo cual debe cumplirse $\frac{\partial f}{\partial x}(\boldsymbol{x_0}) = 0$ y $\frac{\partial f}{\partial y}(\boldsymbol{x_0}) = 0$.

Demostración de la Proposición 2.2.4. Analizaremos únicamente el caso en el cual f alcanza un máximo relativo en $\boldsymbol{x_0}$, ya que el otro es totalmente análogo. Debemos probar que todas las derivadas parciales de f (las cuales existen por ser f diferenciable) se anulan en $\boldsymbol{x_0}$, es decir, que $\frac{\partial f}{\partial x_i}(\boldsymbol{x_0}) = 0$ para todo $1 \leq i \leq n$. Veamos esto para un i fijo y arbitrario.

Escribimos $\boldsymbol{x_0} = (x_0^1, \ldots, x_0^i, \ldots, x_0^n)$, y definimos la función $g\colon \mathbb{R} \to \mathbb{R}$ como

$$g(t) = f(x_0^1, \ldots, x_0^{i-1}, t, x_0^{i+1}, \ldots, x_0^n).$$

Esta función es claramente derivable ya que f es diferenciable. Más aún, por la Regla de la Cadena se tiene para $t = x_0^i$ que:

$$g'(x_0^i) = \begin{pmatrix} \frac{\partial f}{x_1}(\boldsymbol{x_0}) & \cdots & \frac{\partial f}{x_i}(\boldsymbol{x_0}) & \cdots & \frac{\partial f}{x_n}(\boldsymbol{x_0}) \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \frac{\partial f}{x_i}(\boldsymbol{x_0}).$$

Por otro lado, como f alcanza un máximo relativo en $\boldsymbol{x_0}$, entonces existe $\delta > 0$ tal que

$$f(\boldsymbol{x_0}) \geq f(\boldsymbol{x}) \text{ para todo } \boldsymbol{x} \in B(\boldsymbol{x_0}, \delta) \cap U.$$

Note que $f(\boldsymbol{x_0}) = g(x_0^i)$ y que $f(\boldsymbol{x}) = g(x_i)$ para $\boldsymbol{x} = (x_1, \ldots, x_n)$. Entonces, g alcanza un máximo relativo en x_0^i, por lo cual $g'(x_0^i) = 0$. A partir de la igualdad anterior se tiene

$$\frac{\partial f}{x_i}(\boldsymbol{x_0}) = g'(x_0^i) = 0.$$

Como i es arbitrario, se tiene que $\nabla f(\boldsymbol{x_0}) = 0$, por lo cual $df_{\boldsymbol{x_0}} = 0$. $\square$

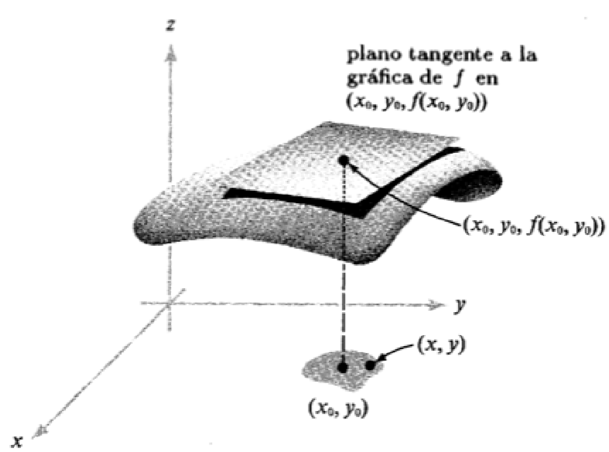

Figura 2.4: Plano tangente perpendicular al eje Z en los puntos donde se alcanza un máximo relativo.
(Imagen tomada de Cálculo Vectorial - Marsden & Tromba)

2.3 Criterio del diferencial de orden 2

Para poder entender cómo clasificar los puntos críticos de un campo escalar en términos del diferencial de orden dos, es necesario recordar el Teorema de Taylor en varias variables.

Teorema 2.3.1 (aproximación por polinomios de Taylor). *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar de clase C^{k+1} (es decir, f posee derivadas parciales continuas hasta orden $k+1$ en cualquier punto de su dominio U), y sea $\boldsymbol{x_0} \in U$. Entonces,*

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) = f(\boldsymbol{x_0}) + df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + \frac{1}{2}d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + \frac{1}{3!}d^3_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + \cdots + \frac{1}{k!}d^k f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + r_k(\boldsymbol{\Delta x}),$$

donde r_k es un campo escalar definido en un entorno de $\boldsymbol{x_0}$ que además cumple la condición

$$\lim_{\boldsymbol{\Delta x} \to \boldsymbol{0}} \frac{r_k(\boldsymbol{\Delta x})}{||\boldsymbol{\Delta x}||^k} = 0.$$

Para la clasificación de puntos estacionarios solamente nos interesará el caso $k = 2$ del teorema anterior. Vamos a enunciarlo a continuación para nuestra comodidad.

Corolario 2.3.2 (aproximación por polinomios de Taylor para funciones de clase C^3). *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar de clase C^3 y $\boldsymbol{x_0} \in U$. Entonces,*

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) = f(\boldsymbol{x_0}) + df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + \frac{1}{2}d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + r_2(\boldsymbol{\Delta x}),$$

donde r_2 es un campo escalar definido en un entorno de $\boldsymbol{x_0}$ que cumple $\lim_{\boldsymbol{\Delta x} \to \boldsymbol{0}} \frac{r_2(\boldsymbol{\Delta x})}{||\boldsymbol{\Delta x}||^2} = 0$.

Anteriormente mencionamos que la parte $f(\boldsymbol{x_0}) + df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x})$ representa una aproximación a f en un entorno de $\boldsymbol{x_0}$ y que corresponde al plano tangente a la gráfica de f que pasa por $(\boldsymbol{x_0}, f(\boldsymbol{x_0}))$. Para el caso en el cual $\boldsymbol{x_0}$ es un punto estacionario de f, vamos a tener que $df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) = 0$, y por el corolario anterior nos va a quedar

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) = f(\boldsymbol{x_0}) + \frac{1}{2}d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + r_2(\boldsymbol{\Delta x})$$

en un entorno de $\boldsymbol{x_0}$. Otra forma de expresar esto es como una aproximación:

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) \approx f(\boldsymbol{x_0}) + \frac{1}{2}d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}).$$

El término $d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x})$ es un polinomio de grado 2 y homogéneo (es decir, todos sus términos tienen grado 2), en variables $\boldsymbol{\Delta x} = (\Delta x_1, \ldots, \Delta x_n)$. Para el caso de dos variables, $d^2_{\boldsymbol{x_0}}(\boldsymbol{\Delta x})$ es de la forma $A(\Delta x)^2 + B(\Delta y)^2 + C\Delta x \Delta y$. Entonces, localmente se tiene que cerca de $\boldsymbol{x_0}$ la gráfica de f se parece a una superficie cuádrica (por ejemplo, un paraboloide elíptico, hiperbólico, hiperboloide). Por cálculos hechos en ejemplos anteriores, para este tipo de superficies es más fácil determinar cuando en un punto estacionario se alcanza un extremo relativo. Además, hay toda una teoría algebraica que respalda esta observación, y tiene que ver con funciones llamadas formas cuadráticas, las cuales se estudian al final del curso de GAL 2.

Definición 2.3.3 (formas cuadráticas)**.** *Una **forma cuadrática** es una función $q\colon \mathbb{R}^n \to \mathbb{R}$ de la forma*

$$q(\boldsymbol{x}) = \boldsymbol{x}^t \cdot A \cdot \boldsymbol{x} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

donde

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

es una matriz simétrica (es decir, $a_{ij} = a_{ji}$).

Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar de clase C^3 y $\boldsymbol{x_0} \in U$. Entonces, $d^2 f_{\boldsymbol{x_0}}\colon \mathbb{R}^n \to \mathbb{R}$ es una forma cuadrática. La matriz que corresponde a $d^2 f_{\boldsymbol{x_0}}$ (en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}$) se conoce como **matriz Hessiana de $\boldsymbol{f}$ en $\boldsymbol{x_0}$** y se define como:

$$H_f(\boldsymbol{x_0}) := \left(D_{ij} f(\boldsymbol{x_0})\right)_{n \times n},$$

donde

$$D_{ij} f(\boldsymbol{x_0}) := \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x_0}).$$

Es decir,

$$H_f(\boldsymbol{x_0}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x_0}) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{x_0}) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x_0}) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2}(\boldsymbol{x_0}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x_0}) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\boldsymbol{x_0}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x_0}) \end{pmatrix}.$$

Se sabe por un resultado de CDIVV que para funciones de clase C^2, las derivadas parciales cruzadas de orden 2 coinciden, es decir, $\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x_0}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x_0})$ para todo $1 \leq i, j \leq n$. Entonces, $H_f(\boldsymbol{x_0})$ es una matriz simétrica.

Se pueden clasificar los puntos estacionarios $\boldsymbol{x_0}$ de un campo escalar f de clase C^3 conociendo qué tipo de forma cuadrática es $d^2 f_{\boldsymbol{x_0}}$. A su vez, esto último se puede saber averiguando los signos de los valores propios de $H_f(\boldsymbol{x_0})$. Recuerde del curso de GAL2 (más específicamente, del Teorema Espectral para matrices simétricas) que toma matriz simétrica es diagonalizable y posee todos sus valores propios en $\mathbb{R}$. Aplicando esto a $H_f(\boldsymbol{x_0})$, se tiene el siguiente resultado, que se conoce como criterio del diferencial de orden 2, y que representa la generalización en varias variables para el criterio de la derivada segunda. Por el momento daremos solo el enunciado y un ejemplo para interpretar la información que arroja. La demostración la veremos después de recordar los tipos de formas canónicas.

Teorema 2.3.4 (criterio del diferencial de orden 2)**.** *Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ un campo escalar de clase C^3 y $\boldsymbol{x_0} \in U$ un punto estacionario de f. Entonces, las siguientes afirmaciones referentes a la matriz Hessiana $H_f(\boldsymbol{x_0})$ se cumplen:*

1. *Si todos los valores propios de $H_f(\boldsymbol{x_0})$ son positivos, entonces f alcanza un mínimo relativo en $\boldsymbol{x_0}$.*

2. *Si todos los valores propios de $H_f(\boldsymbol{x_0})$ son negativos, entonces f tiene un máximo relativo en $\boldsymbol{x_0}$.*

3. *Si $H_f(\boldsymbol{x_0})$ tiene valores propios positivos y negativos, entonces f presenta un punto de silla en $\boldsymbol{x_0}$.*

4. *Si $H_f(\boldsymbol{x_0})$ tiene al menos un valor propio igual a cero, el criterio no decide.*

Veamos con un ejemplo qué significa la información que arroja este criterio.

Ejemplo 2.3.5. *Sea $f\colon \mathbb{R}^3 \to \mathbb{R}$ el campo escalar dado por $f(x,y,z) = x^4 + y^4 + z^4 - 4xyz$. Veamos que en el punto $(1,1,1)$ se alcanza un mínimo relativo.*

Lo primero que se debe hacer, a la hora de querer aplicar el criterio del diferencial de orden 2, es hallar los puntos estacionarios de f. En efecto, por la Proposición 2.2.4 tenemos que los puntos donde f alcanza sus extremos relativos (de haberlos) se encuentran dentro del conjunto de todos los puntos estacionarios de f. Entonces, veamos primero que $(1,1,1)$ anula a todas las derivadas parciales de f. Tenemos de manera general lo siguiente:

$$\frac{\partial f}{\partial x} = 4x^3 - 4yz, \qquad \frac{\partial f}{\partial y} = 4y^3 - 4xz, \qquad \frac{\partial f}{\partial z} = 4z^3 - 4xy.$$

Entonces,

$$\frac{\partial f}{\partial x}(1,1,1) = 4(1)^3 - 4 \cdot 1 \cdot 1 = 0, \qquad \frac{\partial f}{\partial y}(1,1,1) = 0, \qquad \frac{\partial f}{\partial z}(1,1,1) = 0.$$

Para las derivadas parciales de orden 2, tenemos de manera general lo siguiente:

$$
\begin{array}{lll}
\dfrac{\partial^2 f}{\partial x^2} = 12x^2, & \dfrac{\partial^2 f}{\partial x \partial y} = -4z, & \dfrac{\partial^2 f}{\partial z \partial x} = -4y, \\
\dfrac{\partial^2 f}{\partial y \partial x} = -4z, & \dfrac{\partial^2 f}{\partial y^2} = 12y^2, & \dfrac{\partial^2 f}{\partial y \partial z} = -4x, \\
\dfrac{\partial^2 f}{\partial x \partial z} = -4y, & \dfrac{\partial^2 f}{\partial z \partial y} = -4x, & \dfrac{\partial^2 f}{\partial z^2} = 12z^2.
\end{array}
$$

Entonces, la matriz Hessiana de f en $(1,1,1)$ es $H_f(1,1,1) = \begin{pmatrix} 12 & -4 & -4 \\ -4 & 12 & -4 \\ -4 & -4 & 12 \end{pmatrix}$.

Recuerde que para obtener los valores propios de esta matriz se calculan las raíces de su polinomio característico $\chi_{H_f(1,1,1)}(\lambda)$. Recordemos que $\chi_{H_f(1,1,1)}(\lambda)$ se define como $\chi_{H_f(1,1,1)}(\lambda) = \det(H_f(1,1,1) - \lambda I_{3\times 3})$, donde $I_{3\times 3}$ denota la matriz identidad de orden 3:

$$
\chi_{H_f(1,1,1)}(\lambda) = \det \begin{pmatrix} 12-\lambda & -4 & -4 \\ -4 & 12-\lambda & -4 \\ -4 & -4 & 12-\lambda \end{pmatrix} = (16-\lambda)^2(4-\lambda).
$$

La última igualdad se debe a propiedades de los determinantes. Tenemos entonces que 4 y 16 (de multiplicidad algebraica 2) son los valores propios de $H_f(1,1,1)$, y son positivos. Por el criterio del diferencial de orden 2, tenemos que f alcanza un mínimo local en $(1,1,1)$.

Ahora, veamos que significa la conclusión obtenida. De acuerdo con los cálculos hechos, tenemos que f se puede aproximar alrededor del punto $(1,1,1)$ de la siguiente forma:

$$
\begin{aligned}
f((1,1,1) + \boldsymbol{\Delta x}) &\approx f(1,1,1) + df_{(1,1,1)}(\boldsymbol{\Delta}) + \frac{1}{2} d^2 f_{(1,1,1)}(\boldsymbol{\Delta}) \\
&\approx -1 + \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} 12 & -4 & -4 \\ -4 & 12 & -4 \\ -4 & -4 & 12 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \\
&\approx \begin{pmatrix} \Delta x & \Delta y & \Delta z \end{pmatrix} \begin{pmatrix} 6\Delta x - 2\Delta y - 2\Delta z \\ -2\Delta x + 6\Delta y - 2\Delta z \\ -2\Delta x - 2\Delta y + 6\Delta z \end{pmatrix} - 1 \\
&\approx 6(\Delta x)^2 + 6(\Delta y)^2 + 6(\Delta z)^2 - 4\Delta x \Delta z - 4\Delta y \Delta z \\
&\approx 2(\Delta x - \Delta y)^2 + 2(\Delta x - \Delta z)^2 + 2(\Delta y - \Delta z)^2 + 2(\Delta x)^2 + 2(\Delta y)^2 + 2(\Delta z)^2 - 1
\end{aligned}
$$

Es decir, que localmente, f se aproxima al polinomio de grado 2 anterior, el cual a su vez toma siempre valores mayores o iguales a -1. Luego, en los puntos donde este polinomio vale -1 (en particular en el punto $(1,1,1)$) se alcanzan mínimos locales.

Recordemos ahora un poco de teoría de formas cuadráticas para poder abordar la demostración del Criterio del Diferencial de Orden 2.

Definición 2.3.6. *Sea $q\colon \mathbb{R}^n \to \mathbb{R}$ una forma cuadrática. Entonces, q se dice:*

1. ***definida positiva*** *si $q(\boldsymbol{x}) > 0$ para todo $\boldsymbol{x} \in \mathbb{R}^n - \{\mathbf{0}\}$;*

2. ***definida negativa*** *si $q(\boldsymbol{x}) < 0$ para todo $\boldsymbol{x} \in \mathbb{R}^n - \{\mathbf{0}\}$;*

3. ***indefinida*** *si existen $\boldsymbol{x_1}, \boldsymbol{x_2} \in \mathbb{R}^n$ tales que $q(\boldsymbol{x_1}) > 0$ y $q(\boldsymbol{x_2}) < 0$.*

Teorema 2.3.7 (teorema de clasificación de formas cuadráticas)**.** *Sea $q\colon \mathbb{R}^n \to \mathbb{R}$ una forma cuadrática de la forma $\boldsymbol{x} \mapsto \boldsymbol{x}^t A \boldsymbol{x}$ ($A \in M_{n\times n}(\mathbb{R})$ simétrica). Entonces, las siguientes equivalencias se cumplen:*

1. *q es definida positiva si, y solo si, todos los valores propios de A son positivos.*

2. *q es definida negativa si, y solo si, todos los valores propios de A son negativos.*

3. *q es indefinida si, y solo si, A tiene valores propios positivos y negativos.*

Podemos notar entonces lo siguiente, para un campo escalar $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase C^3 y con punto estacionario en $\boldsymbol{x_0} \in U$:

1. $d^2 f_{\boldsymbol{x_0}}$ definida positiva $\Longrightarrow f$ alcanza un mínimo relativo en $\boldsymbol{x_0}$.

2. $d^2 f_{\boldsymbol{x_0}}$ definida negativa $\Longrightarrow f$ alcanza un máximo relativo en $\boldsymbol{x_0}$.

3. $d^2 f_{\boldsymbol{x_0}}$ indefinida $\Longrightarrow f$ alcanza un punto de silla en $\boldsymbol{x_0}$.

Estamos listos para dar la prueba del Criterio del Diferencial de Orden 2.

Lema 2.3.8. *Sea $q\colon \mathbb{R}^n \to \mathbb{R}$ una forma cuadrática definida positiva. Entonces existe una constante $m > 0$ tal que $q(\boldsymbol{x}) \geq m||\boldsymbol{x}||^2$ para todo $\boldsymbol{x} \in \mathbb{R}^n$.*

Demostración: La idea para hallar tal m es aplicar el Teorema de Weierstrass a cierto campo escalar continuo sobre cierto dominio compacto. Tal campo es la forma q (es continua por ser un polinomio). El dominio compacto a escoger es la esfera unitaria S^{n-1} (de radio 1 y centro $\mathbf{0}$). Tenemos entonces que existe $m \in \mathbb{R}$ y $\boldsymbol{p} \in S^{n-1}$ tal que $q(\boldsymbol{p}) = m$ y $q(\boldsymbol{x}) \geq m$ para todo $\boldsymbol{x} \in S^{n-1}$. Por otro lado, al ser q definida positiva, $m = q(\boldsymbol{p}) > 0$ ya que $\boldsymbol{p} \neq \mathbf{0}$.

Ahora consideremos cualquier $\boldsymbol{x} \in \mathbb{R}^n$ no nulo. Entonces, $\boldsymbol{x}/||\boldsymbol{x}|| \in S^{n-1}$, por lo cual

$$q\left(\frac{\boldsymbol{x}}{||\boldsymbol{x}||}\right) \geq m.$$

Por otro lado,

$$q\left(\frac{\boldsymbol{x}}{||\boldsymbol{x}||}\right) = \left(\frac{\boldsymbol{x}}{||\boldsymbol{x}||}\right)^t \cdot A \cdot \frac{\boldsymbol{x}}{||\boldsymbol{x}||} = \frac{1}{||\boldsymbol{x}||^2}\boldsymbol{x}^t \cdot A \cdot \boldsymbol{x} = \frac{q(\boldsymbol{x})}{||\boldsymbol{x}||^2}.$$

Se tiene entonces que

$$\frac{q(\boldsymbol{x})}{||\boldsymbol{x}||^2} \geq m,$$

es decir,

$$q(\boldsymbol{x}) \geq m||\boldsymbol{x}||^2$$

para todo $\boldsymbol{x} \in \mathbb{R}^n$ no nulo. Finalmente, es claro que $\mathbf{0}$ también satisface la desigualdad anterior. Por lo tanto, se tiene el resultado. $\square$

Demostración del Teorema 2.3.4. Empezamos aplicando el teorema de Taylor respecto al punto estacionario $\boldsymbol{x_0} \in U$. Tenemos:

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) - f(\boldsymbol{x_0}) = \frac{1}{2} d^2 f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) + r_2(\boldsymbol{\Delta x}),$$

donde $\boldsymbol{\Delta x} = \boldsymbol{x} - \boldsymbol{x_0}$, $\lim_{\boldsymbol{\Delta} \to \mathbf{0}} r_2(\boldsymbol{\Delta x}) = 0$, $\lim_{\boldsymbol{\Delta x} \to \mathbf{0}} \frac{r_2(\boldsymbol{\Delta x})}{||\boldsymbol{\Delta x}||^2} = 0$, y $d^2 f_{\boldsymbol{x_0}} \colon \mathbb{R}^n \to \mathbb{R}$ es la forma cuadrática representada por la matriz simétrica $H_f(\boldsymbol{x_0})$. Probemos primero el caso donde $H_f(\boldsymbol{x_0})$ tiene valores propios positivos (es decir, si $d^2 f_{\boldsymbol{x_0}}$ es definida positiva). Notamos de la igualdad anterior que f va a alcanzar un mínimo relativo en $\boldsymbol{x_0}$ si $f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) - f(\boldsymbol{x_0}) \geq 0$, con $\boldsymbol{x}$ en un entorno de $\boldsymbol{x_0}$.

Por el lema anterior, existe $m > 0$ tal que $d^2 f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x}) \geq m||\boldsymbol{\Delta x}||^2$. Luego,

$$f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) - f(\boldsymbol{x_0}) \geq \frac{m}{2}||\boldsymbol{\Delta x}||^2 + r_2(\boldsymbol{\Delta x}).$$

Ahora bien, como $\lim_{\boldsymbol{\Delta x} \to \mathbf{0}} \frac{r_2(\boldsymbol{\Delta x})}{||\boldsymbol{\Delta x}||^2} = 0$, para $\varepsilon = m/2$ existe $\delta > 0$ tal que $\left|\frac{r_2(\boldsymbol{\Delta x})}{||\boldsymbol{\Delta x}||^2}\right| < \frac{m}{2}$ para todo $\boldsymbol{\Delta x} \in B(\mathbf{0}, \delta)$. Lo anterior implica que para $\boldsymbol{\Delta x} \in B(\mathbf{0}, \delta)$ se tiene $r_2(\boldsymbol{\Delta x}) > -\frac{m}{2}||\boldsymbol{\Delta x}||^2$. Note que decir $\boldsymbol{\Delta x} \in B(\mathbf{0}, \delta)$ es equivalente a $\boldsymbol{x} \in B(\boldsymbol{x_0}, \delta)$. Entonces, para $\boldsymbol{x} \in B(\boldsymbol{x_0}, \delta) \cap U$ se tiene

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x_0}) = f(\boldsymbol{x_0} + \boldsymbol{\Delta x}) - f(\boldsymbol{x_0}) &\geq \frac{m}{2}||\boldsymbol{x} - \boldsymbol{x_0}||^2 + r_2(\boldsymbol{x} - \boldsymbol{x_0}) \\ &> \frac{m}{2}||\boldsymbol{x} - \boldsymbol{x_0}||^2 - \frac{m}{2}||\boldsymbol{x} - \boldsymbol{x_0}||^2 = 0. \end{aligned}$$

Por lo tanto, $f(\boldsymbol{x}) > f(\boldsymbol{x_0})$ para $\boldsymbol{x} \in B(\boldsymbol{x_0}, \varepsilon) \cap U$, es decir, f alcanza un mínimo local en $\boldsymbol{x_0}$.

El caso donde $d^2 f_{\boldsymbol{x_0}}$ es definida negativa se sigue del caso anterior aplicado a $g = -f$.

Finalmente, falta probar el caso donde $d^2 f_{\boldsymbol{x_0}}$ es indefinida, es decir, existen $\boldsymbol{x_1}, \boldsymbol{x_2} \in \mathbb{R}^n$ tales que $d^2 f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x_1}) < 0$ y $d^2 f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x_2}) > 0$. Tenemos que encontrar un par de puntos $\boldsymbol{p_1}, \boldsymbol{p_2} \in U$ tales que $f(\boldsymbol{p_2}) > f(\boldsymbol{x_0}) > f(\boldsymbol{p_1})$. Estos puntos van a encontrarse a lo largo de rectas que pasen por $\boldsymbol{x_0}$, por lo que podemos considerar la función $g_{\boldsymbol{\Delta x}} \colon I \subseteq \mathbb{R} \to \mathbb{R}$, con $\boldsymbol{\Delta x}$ fijo, dada por

$$g_{\boldsymbol{\Delta x}}(t) = f(\boldsymbol{x_0} + t\boldsymbol{\Delta x}),$$

para todo $t \in I$, donde I es un intervalo que contiene a 0 y tal que $\boldsymbol{x_0} + t\boldsymbol{\Delta x} \in U$. Esta función $g_{\boldsymbol{\Delta x}}$ es claramente dos veces derivable por ser f de dos veces diferenciable y por la regla de la cadena. Más aún, se tiene que $g'_{\boldsymbol{\Delta x}}(t) = df_{\boldsymbol{x_0}+t\boldsymbol{\Delta x}}(\boldsymbol{\Delta x})$ y $g''_{\boldsymbol{\Delta x}}(t) = d^2 f_{\boldsymbol{x_0}+t\boldsymbol{\Delta x}}(\boldsymbol{\Delta x})$. En particular, para $\boldsymbol{\Delta x} = \boldsymbol{\Delta x_1}$, vamos a tener que $g'_{\boldsymbol{\Delta x_1}}(0) = df_{\boldsymbol{x_0}}(\boldsymbol{\Delta x_1}) = 0$ (por ser $\boldsymbol{x_0}$ un punto estacionario) y $g''_{\boldsymbol{\Delta x_1}}(0) = d^2 f_{\boldsymbol{x_0}}(\boldsymbol{\Delta x_1}) < 0$, es decir, que $g_{\boldsymbol{\Delta x_1}}$ alcanza un máximo relativo en $t = 0$, es decir,

$$f(\boldsymbol{x_0} + t\boldsymbol{\Delta x_1}) = g_{\boldsymbol{\Delta x_1}}(t) \leq g_{\boldsymbol{\Delta x_1}}(0) = f(\boldsymbol{x_0})$$

para t en un entorno de 0. Entonces existe t_1 tal que $f(\boldsymbol{x_0} + t_1\boldsymbol{\Delta x_1}) \leq f(\boldsymbol{x_0})$. De manera similar, usando $\boldsymbol{\Delta x_2}$, existe t_2 tal que $f(\boldsymbol{x_0} + t_2\boldsymbol{\Delta x_2}) \geq f(\boldsymbol{x_0})$. Se sigue entonces el resultado en este caso tomando $\boldsymbol{p_1} = \boldsymbol{x_0} + t_1\boldsymbol{\Delta x_1}$ y $\boldsymbol{p_2} = \boldsymbol{x_0} + t_2\boldsymbol{\Delta x_2}$. $\square$

Este criterio tiene un enunciado equivalente un poco más manejable, a la hora de hacer cálculos, para campos escalares de dos variables. En efecto, supongamos que tenemos un campo escalar $f \colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ donde $\boldsymbol{x_0} \in U$ es un punto estacionario. Consideremos el diferencial de orden 2 de f en $\boldsymbol{x_0}$, es decir, la forma cuadrática $d^2 f_{\boldsymbol{x_0}} \colon \mathbb{R}^2 \to \mathbb{R}$ representada por la matriz simétrica Hessiana

$$H_f(\boldsymbol{x_0}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\boldsymbol{x_0}) & \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{x_0}) \\ \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{x_0}) & \frac{\partial^2 f}{\partial y^2}(\boldsymbol{x_0}) \end{pmatrix}.$$

1. Supongamos que $d^2 f_{\boldsymbol{x_0}}$ es definida positiva, es decir, que $H_f(\boldsymbol{x_0})$ tiene todos sus valores propios positivos. Entonces, al ser $H_f(\boldsymbol{x_0})$ diagonalizable, tenemos que $H_f(\boldsymbol{x_0})$ es semejante a la matriz diagonal formada por sus valores propios, a los que llamaremos λ_1 y λ_2. Entonces,

   $$H_f(\boldsymbol{x_0}) \sim \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

   Como la traza y determinante de matrices semejantes coinciden, vamos a tener que

   $$\det(H_f(\boldsymbol{x_0})) = \lambda_1\lambda_2 \qquad \text{y} \qquad \operatorname{tr}(H_f(\boldsymbol{x_0})) = \lambda_1 + \lambda_2.$$

   Como $\lambda_1, \lambda_2 > 0$, tenemos que $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) > 0$.

Ahora supongamos que $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) > 0$, es decir, $\lambda_1\lambda_2 > 0$ y $\lambda_1 + \lambda_2 > 0$. La primera desigualdad implica que λ_1 y λ_2 son ambos positivos o negativos. El segundo escenario no puede ocurrir ya que $\lambda_1 + \lambda_2 > 0$. Por lo tanto, $\lambda_1, \lambda_2 > 0$. Entonces, $d^2 f_{\boldsymbol{x_0}}$ es definida positiva si, y solo si, $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) > 0$.

2. Análogamente, $d^2 f_{\boldsymbol{x_0}}$ definida negativa equivale a $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) < 0$.

3. Ahora supongamos que $d^2 f_{\boldsymbol{x_0}}$ es indefinida. Entonces, $\lambda_1 > 0$ y $\lambda_2 < 0$ (o $\lambda_1 < 0$ y $\lambda_2 > 0$). En cualquier caso, se tiene que $\det(H_f(\boldsymbol{x_0})) < 0$. Recíprocamente, $\lambda_1\lambda_2 = \det(H_f(\boldsymbol{x_0})) < 0$ implica que λ_1 y λ_2 tienen signos diferentes. Por lo tanto, $d^2 f_{\boldsymbol{x_0}}$ es indefinida si, y solo si, $\det(H_f(\boldsymbol{x_0})) < 0$.

De lo anterior se tiene el siguiente resultado.

Corolario 2.3.9 (criterio del diferencial de orden 2 - para campos escalares de dos variables)**.** *Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ un campo escalar de clase C^3 y $\boldsymbol{x_0} \in U$ un punto estacionario de f. Entonces, las siguientes afirmaciones referentes a la matriz Hessiana $H_f(\boldsymbol{x_0})$ se cumplen:*

1. *Si $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) > 0$, entonces f alcanza un mínimo relativo en $\boldsymbol{x_0}$.*

2. *Si $\det(H_f(\boldsymbol{x_0})) > 0$ y $\operatorname{tr}(H_f(\boldsymbol{x_0})) < 0$, entonces f alcanza un máximo relativo en $\boldsymbol{x_0}$.*

3. *Si $\det(H_f(\boldsymbol{x_0})) < 0$, entonces f presenta un punto de silla en $\boldsymbol{x_0}$.*

4. *Si $\det(H_f(\boldsymbol{x_0})) = 0$, el criterio no decide.*

Ejemplo 2.3.10. *Hallar y clasificar los puntos estacionarios del campo escalar $f\colon \mathbb{R}^2 \to \mathbb{R}$ dado por*

$$f(x,y) = (x^2+y^2)e^{-(x^2+y^2)}.$$

Lo primero a hacer es calcular lar derivadas parciales de orden 1:

$$\begin{aligned}\frac{\partial f}{\partial x} &= (2x - 2x^3 - 2xy^2)e^{-(x^2+y^2)} = 2x(1-x^2-y^2)e^{-(x^2+y^2)},\\ \frac{\partial f}{\partial y} &= (2y - 2y^3 - 2x^2y)e^{-(x^2+y^2)} = 2y(1-x^2-y^2)e^{-(x^2+y^2)}.\end{aligned}$$

Tenemos entonces que los puntos estacionarios están dados por $(0,0)$ y por todos los puntos ubicados en la circunferencia S^1 de centro $\mathbf{0}$ y radio 1.

Ahora, para clasificar estos puntos críticos, necesitamos las derivadas parciales de orden 2:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= 2(1-5x^2-y^2+2x^2y^2+2x^4)e^{-(x^2+y^2)},\\
\frac{\partial^2 f}{\partial y^2} &= 2(1-5y^2-x^2+2x^2y^2+2y^4)e^{-(x^2+y^2)},\\
\frac{\partial^2 f}{\partial x \partial y} &= -4x(y-x+x^3+xy^2)e^{-(x^2+y^2)}.
\end{aligned}$$

Tenemos así que la matriz Hessiana en cualquier punto viene dada por:

$$\begin{aligned}
H_f(x,y) &= \begin{pmatrix} 2(1-5x^2-y^2+2x^2y^2+2x^4)e^{-(x^2+y^2)} & -4x(y-x+x^3+xy^2)e^{-(x^2+y^2)} \\ -4x(y-x+x^3+xy^2)e^{-(x^2+y^2)} & 2(1-5y^2-x^2+2x^2y^2+2y^4)e^{-(x^2+y^2)} \end{pmatrix}\\
&= 2\begin{pmatrix} 1-5x^2-y^2+2x^2y^2+2x^4 & -2x(y-x+x^3+xy^2) \\ -2x(y-x+x^3+xy^2) & 1-5y^2-x^2+2x^2y^2+2y^4 \end{pmatrix} e^{-(x^2+y^2)}.
\end{aligned}$$

Analicemos primero el punto estacionario $(0,0)$. *Del cálculo anterior tenemos*

$$H_f(0,0) = 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} e^0 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Luego, $\det(H_f(0,0))4 > 0$ *y* $\mathrm{tr}(H_f(0,0)) = 4 > 0$. *Por el criterio del diferencial de orden 2, tenemos que* f *alcanza un mínimo relativo en* $(0,0)$. *Podemos decir más sobre* $(0,0)$. *Tal mínimo* $f(0,0) = 0$ *es de hecho absoluto, pues* $f(x,y) \geq 0$ *claramente para todo* $(x,y) \in \mathbb{R}^2$.

Ahora analicemos el conjunto restante de puntos estacionarios S^1. *Sea* $(x_0,y_0) \in S^1$, *es decir,* $x_0^2+y_0^2 = 1$. *Luego,*

$$\begin{aligned}
H_f(x_0,y_0) &= 2\begin{pmatrix} 1-5x_0^2-y_0^2+2x_0^2y_0^2+2x_0^4 & -2x_0(y_0-x_0+x_0^3+x_0y_0^2) \\ -2x_0(y_0-x_0+x_0^3+x_0y_0^2) & 1-5y_0^2-x_0^2+2x_0^2y_0^2+2y_0^4 \end{pmatrix} e^{-(x_0^2+y_0^2)}\\
&= 2\begin{pmatrix} 1-4x_0^2-(x_0^2+y_0^2)+2x_0^2(y_0^2+x_0^2) & -2x_0(y_0-x_0+x_0(x_0^2+y_0^2)) \\ -2x_0(y_0-x_0+x_0(x_0^2+y_0^2)) & 1-4y_0^2-(y_0^2+x_0^2)+2y_0^2(x_0^2+y_0^2) \end{pmatrix} e^{-(x_0^2+y_0^2)}\\
&= 2\begin{pmatrix} 1-4x_0^2-1+2x_0^2 & -2x_0(y_0-x_0+x_0) \\ -2x_0(y_0-x_0+x_0) & 1-4y_0^2-1+2y_0^2 \end{pmatrix} e^{-1}\\
&= \frac{2}{e}\begin{pmatrix} -2x_0^2 & -2x_0y_0 \\ -2x_0y_0 & -2y_0^2 \end{pmatrix} = -\frac{4}{e}\begin{pmatrix} x_0^2 & x_0y_0 \\ x_0y_0 & y_0^2 \end{pmatrix}.
\end{aligned}$$

Tenemos entonces que

$$\det(H_f(x_0,y_0)) = \frac{16}{e^2}\det\begin{pmatrix} x_0^2 & x_0y_0 \\ x_0y_0 & y_0^2 \end{pmatrix} = \frac{16}{e^2}(x_0^2y_0^2 - x_0^2y_0^2) = 0.$$

En este caso tenemos que el criterio del diferencial de orden 2 no decide. Tenemos entonces que buscar otra forma de saber qué tipo de punto estacionario es (x_0, y_0)*. Esto usualmente depende de propiedades intrínsecas de la fórmula que define a la función* f *o de su dominio, por lo que no hay un método general sobre cómo proceder en estos casos.*

Lo primero a notar es que $f(x_0, y_0) = e^{-1}$*. Por otro lado, a* $f(x, y) = (x^2 + y^2)e^{-(x^2+y^2)}$ *podemos verla como una función de una variable haciendo* $t = x^2 + y^2 \geq 0$*. Definimos así una nueva función* $g\colon [0, +\infty) \to \mathbb{R}$ *dada por* $g(t) = te^{-t}$*. Notamos que* $g(1) = f(x_0, y_0)$ *y que* $te^{-t} \leq e^{-1}$*, o equivalentemente,* $t \leq e^{t-1}$*. Lo último se puede ver fácilmente graficando la identidad y la función* $t \mapsto e^{t-1}$*. Por lo tanto, devolviendo el cambio* $t = x^2 + y^2$*, tenemos que* $f(x, y) \leq f(x_0, y_0)$*, es decir,* e^{-1} *es el máximo absoluto de* f *alcanzado en cualquier punto de la circunferencia* S^1*.*

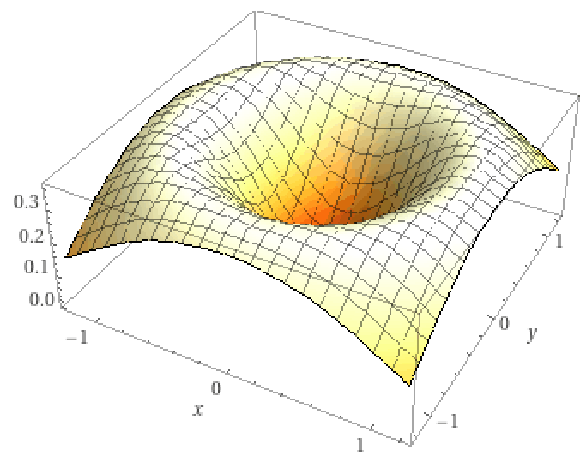

Figura 2.5: Gráfica de f. `(Wolfram Alpha)`

Escrito en LaTeX por **Marco A. Pérez**.

Material consultado:
- Cálculo Vectorial, de Marsden y Tromba.
- Calculus, Vol. 2, de Apostol.

Última actualización: `26 de Agosto de 2020`.