March 5, 2020

2. OPTIMIZACIÓN SIN RESTRICCIONES

En toda esta sección D denota un **subconjunto abierto** de $\mathbb{R}^n$.

1. Condiciones Necesarias de Primer Orden

Proposición 1.1. Sea $f : D \to \mathbb{R}$ diferenciable. Si $p \in D$ es un máximo o un mínimo local de f en D, entonces

$$\nabla f(p) = 0$$

Demostración Fijemos $i = 1 \ldots, n$ y consideremos la curva

$$g(t) = f(p + te_i)$$

donde $\{e_1 \ldots, e_n\}$ es la base canónica de $\mathbb{R}^n$. Observamos que g es una función diferenciable de 1-variable que tiene un máximo local en $t_0 = 0$. Así,

$$g'(0) = 0$$

Pero,

$$g'(0) = \frac{d}{dt}\bigg|_{t=0} f(p + te_i) = \lim_{x \to 0} \frac{f(p + te_i - f(p)}{t} = \frac{\partial f}{\partial x_i}(p)$$

Definición 1.2. Sea $f : D \to \mathbb{R}$ se dice que $p \in D$ es un **punto crítico** si f no es diferenciable en p o si

$$\nabla f(p) = 0$$

Observación 1.3. Si p es un extremo local de f, entonces p es un punto crítico de f.

Definición 1.4. Si $\nabla f(p) = 0$, pero p no es un extremo local de f, entonces p es un **punto de silla**.

2. Condiciones Necesarias de Segundo Orden

Proposición 2.1. Sea $f : D \to \mathbb{R}$ de clase $C^2(D)$. Dado un punto $p \in D$.

(1) Si p es un máximo local de f en D, entonces la matriz Hessiana de $\mathrm{H}\, f(p)$ es semidefinida negativa o definida negativa.
(2) Si p es un mínimo local de f en D, entonces la matriz Hessiana de $\mathrm{H}\, f(p)$ es semidefinida positiva o definida positiva.

3. Condiciones Suficientes de Segundo Orden

Proposición 3.1. Sea $f : D \to \mathbb{R}$ de clase $C^2(D)$. Sea $p \in D$ y supongamos que

$$\nabla f(p) = 0.$$

(1) Si $\mathrm{H}\, f(p)$ es definida negativa , entonces p es un máximo local (estricto) de f.
(2) Si $\mathrm{H}\, f(p)$ es definida positiva, entonces p es un mínimo local (estricto) de f.
(3) Si $\mathrm{H}\, f(p)$ es indefinida, entonces p es un punto de silla.

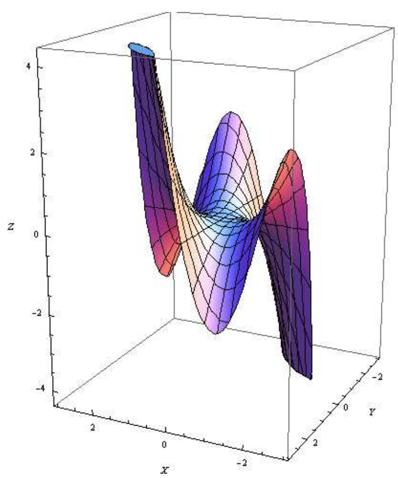
Ejemplo 3.2. Consideremos la función,

$$f(x,y) = x^2y + y^2x$$

Entonces, $\nabla f(x,y) = (2xy + y^2, 2xy + x^2)$ y el único punto crítico es $(0,0)$. Para determinar si es un máximo, mínimo o punto de silla calculamos la matriz Hessiana,

$$\mathrm{H}\, f(0,0) = \begin{pmatrix} 2y & 2x+2y \\ 2x+2y & 2x \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Las derivadas de segundo orden no aportan ninguna información. Pero dado que $f(x,x) = 2x^3$, entonces $(0,0)$ es un punto de silla. La gráfica de f es la siguiente



Ejemplo 3.3. Consideremos la función,

$$f(x,y) = (x-1)^4 + (y-1)^2$$

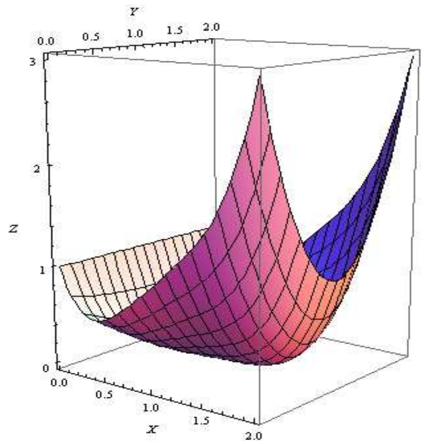
Entonces,

$$\nabla f(x,y) = (4(x-1)^3, 2(y-1))$$

y el único punto crítico es $(1,1)$. Para determinar si es un máximo, mínimo o punto de silla calculamos la matriz Hessiana,

$$\mathrm{H}\, f(1,1) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

ComoH $f(1,1)$ es semidefinido positivo las condiciones de segundo orden no aportan ninguna información. Pero dado que $f(x,y) \geq 0 = f(1,1)$, entonces $(1,1)$ es un mínimo global. La gráfica de f es la siguiente



Ejemplo 3.4. Consideremos la función,

$$f(x,y) = (x-1)^3 + y^2$$

El gradiente es

$$\nabla f(x,y) = (3(x-1)^2, 2y)$$

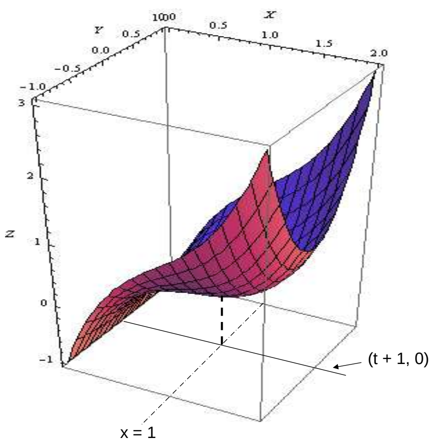
y existe un único punto crítico en $(1,0)$. Para clasificarlo, calculamos la matriz Hessiana

$$\mathrm{H}\, f(1,0) = \begin{pmatrix} 6(x-1), & 0 \\ 0 & 2 \end{pmatrix}\Bigg|_{x=1 y=0} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

como es semidefinida positiva las condiciones de segundo orden no aportan ninguna información. Pero dado que,

$$f(1+t,0) = t^3 = \begin{cases} > 0 \text{ si } t > 0 \\ < 0 \text{ si } t < 0 \end{cases}$$

vemos que $(1,0)$ es un punto de silla. La gráfica de f es la siguiente



Ejemplo 3.5. Consideremos la función,

$$f(x, y) = x^2 + y^2(x + 1)^3$$

El gradiente es

$$\nabla f(x, y) = \big(2x + 3y^2(x + 1)^2, 2y(x + 1)^3\big)$$

y existe un único punto crítico en $(0, 0)$. Para clasificarlo, calculamos la matriz Hessiana,

$$\operatorname{H} f(0, 0) = \begin{pmatrix} 2 + 6y^2(x + 1) & 6y(x + 1)^2 \\ 6y(x + 1)^2 & 2(x + 1)^3 \end{pmatrix}\Bigg|_{x=y=0} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

que es definida positiva. Así, $(0, 0)$ es un mínimo local (estricto). Pero no es mínimo global, porque $f(-2, y) = 4 - y^2$ puede hacerse arbitrariamente pequeño al tomar valores grandes de y.

Observación 3.6 (Una justificación intuitiva de las condiciones de segundo orden). Recordemos que el polinomio de Taylor de orden 2 de f en el punto p es

$$P_2(x) = f(p) + \nabla f(p) \cdot (x - p) + \frac{1}{2}(x - p) \operatorname{H} f(p)(x - p)$$

Recordemos también que si f es de clase C^2 entonces

$$\lim_{x \to 0} \frac{R_2(x)}{\|x - p\|^2} = 0$$

donde

$$R_2(x) = f(x) - P_2(x)$$

es el error cometido al aproximar la función f por el polinomio de Taylor de orden 2. Supongamos ahora que p es un punto crítico de f y, por tanto $\nabla f(p) = 0$. Entonces

$$f(x) - f(p) = \frac{1}{2}(x - p) \operatorname{H} f(p)(x - p) + R_2(x)$$

y para x cercano a p el término $R_2(x)$ es ‘pequeño’. Por tanto si, por ejemplo, sabemos que el término

$$(x - p) \operatorname{H} f(p)(x - p) > 0$$

entonces $f(x) - f(p) > 0$ para todo $x \neq p$ 'suficientemente cercano' a p y el punto p debería ser un mínimo local. Pero la condición $(x-p)\,\mathrm{H}\,f(p)(x-p) > 0$ para todo $x \neq p$ se verifica si H es definido positivo.