Forma matricial de los polinomios de Taylor

Sea $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ abierto, $A = (a_1, a_2, \ldots, a_n) \in D$

- Si $f \in \mathcal{C}^1(D)$, el polinomio de Taylor de orden 1 centrado en A tiene la expresión

$$P_1(X) = f(A) + \nabla f(A) \cdot (X - A)$$

donde $X = (x_1, x_2, \ldots, x_n)$ y $X - A = (x_1 - a_1, \ldots, x_n - a_n)$.

Si llamamos $\vec{h} = X - A$, entonces $X = A + \vec{h}$ y

$$P_1(A + \vec{h}) = f(A) + \nabla f(A) \cdot \vec{h}$$

Como P_1 es la aproximación lineal de f en A, por lo que vimos en la parte de diferenciabilidad, tenemos que

$$f(A+\vec{h}) = P_1(A+\vec{h}) + \vec{e}_1(\vec{h}) \quad \text{con } e_1(\vec{h}) \Big/ \lim_{\vec{h} \to \vec{0}} \frac{\vec{e}_1(\vec{h})}{\|\vec{h}\|} = 0,$$

es decir, $\|\vec{e}_1(\vec{h})\| << \|\vec{h}\|$ si $\|\vec{h}\|$ es pequeña.

De allí que $f(X) = f(A+\vec{h}) \approx P_1(A+\vec{h}) = P_1(X)$ si $\|\vec{h}\|$ es pequeña

- Si $f \in \mathcal{C}^2(D)$, se puede probar que el polinomio de Taylor de orden 2 tiene la expresión:

transpuesta

$$P_2(X) = f(A) + \nabla f(A) \cdot (X - A) + \frac{1}{2} (X-A)^T \, Hf(A) \, (X - A)$$

donde ahora pensamos a $X - A$ como un vector columna:

$$X - A = \begin{bmatrix} x_1 - a_1 \\ x_2 - a_2 \\ \vdots \\ x_n - a_n \end{bmatrix}, \qquad (X-A)^T = \begin{bmatrix} x_1 - a_1 & x_2 - a_2 & \cdots & x_n - a_n \end{bmatrix}$$

y $Hf(A)$ es la matriz Hessiana de f en A, que está definida como:

$$Hf(A) = \underbrace{D(\nabla f)(A)}_{\substack{\text{Jacobiana} \\ \text{de } \nabla f}} = \begin{bmatrix} f''_{x_1x_1}(A) & f''_{x_1x_2}(A) & \cdots & f''_{x_1x_2}(A) \\ f''_{x_2x_1}(A) & f''_{x_2x_2}(A) & \cdots & f''_{x_2x_n}(A) \\ \vdots & & & \\ f''_{x_nx_1}(A) & f''_{x_nx_2}(A) & \cdots & f''_{x_nx_n}(A) \end{bmatrix}$$

Como $f \in \mathcal{C}^2(D)$, $f''_{x_i x_j}(A) = f''_{x_j x_i}(A)$ $\forall i \neq j$, con lo cual $Hf(A)$ es una matriz simétrica.

Si $\vec{h} = X - A$, entonces

$$P_2(A+\vec{h}) = f(A) + \nabla f(A) \cdot \vec{h} + \frac{1}{2} \vec{h}^T Hf(A) \vec{h}$$

Se puede probar que

$$f(A+\vec{h}) = P_2(A+\vec{h}) + \vec{e}_2(\vec{h}) \quad y \quad \lim_{\vec{h} \to \vec{0}} \frac{\vec{e}_2(\vec{h})}{\|\vec{h}\|^2} = 0$$

Entonces $\|\vec{e}_2(\vec{h})\| << \|\vec{h}\|^2$ si $\|\vec{h}\|$ es pequeña

Esto dice que el error que se comete al aproximar con P_2, que es $\vec{e}_2$, es más pequeño que el que tenemos al aproximar con P_1, que es $\vec{e}_1$, pues $\|\vec{e}_2(\vec{h})\| << \|\vec{h}\|^2$ y $\|\vec{e}_1(\vec{h})\| << \|\vec{h}\|$ y $\|\vec{h}\|^2 << \|\vec{h}\|$ cuando $\|\vec{h}\|$ es pequeña (si $\|\vec{h}\| = 0.1$, $\|\vec{h}\|^2 = 0.01$)

Ejemplo: $f(x,y) = 2x^2 + e^x y$, $A = (0,2)$. Calcular los polinomios de Taylor de orden 1 y 2 usando las expresiones matriciales.

$f \in \mathcal{C}^\infty(\mathbb{R}^2)$, por lo tanto existen los polinomios pedidos.

1) $f(0,2) = 2$

2) $f'_x(0,2) = 4x + e^x y \big|_{(0,2)} = 2$, $f'_y(0,2) = e^x \big|_{(0,2)} = 1 \Rightarrow \nabla f(0,2) = (2,1)$

3) $f''_{xx}(0,2) = 4 + e^x y \big|_{(0,2)} = 6$, $f''_{xy}(0,2) = e^x \big|_{(0,2)} = 1$, $f''_{yy}(0,2) = 0$

$$\Rightarrow \quad Hf(0,2) = \begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix}$$

Entonces:

$$P_1(x,y) = \underbrace{2}_{f(0,2)} + \underbrace{(2,1)}_{\nabla f(0,2)} \cdot (x, y-2) = 2 + 2x + y - 2 = 2x + y$$

y

$$P_2(x,y) = P_1(x,y) + \frac{1}{2} \begin{bmatrix} x & y-2 \end{bmatrix} \underbrace{\begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix}}_{Hf(0,2)} \begin{bmatrix} x \\ y-2 \end{bmatrix}$$

Como

$$\frac{1}{2} \begin{bmatrix} x & y-2 \end{bmatrix} \begin{pmatrix} 6 & 1 \\ 1 & 0 \end{pmatrix} \begin{bmatrix} x \\ y-2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x & y-2 \end{bmatrix} \begin{bmatrix} 6x + y - 2 \\ \ldots \end{bmatrix}$$

como

$$\frac{1}{2}\begin{bmatrix}x & y-2\end{bmatrix}\begin{pmatrix}6 & 1\\ 1 & 0\end{pmatrix}\begin{bmatrix}x\\ y-2\end{bmatrix} = \frac{1}{2}\begin{bmatrix}x & y-2\end{bmatrix}\begin{bmatrix}6x+y-2\\ x\end{bmatrix}$$

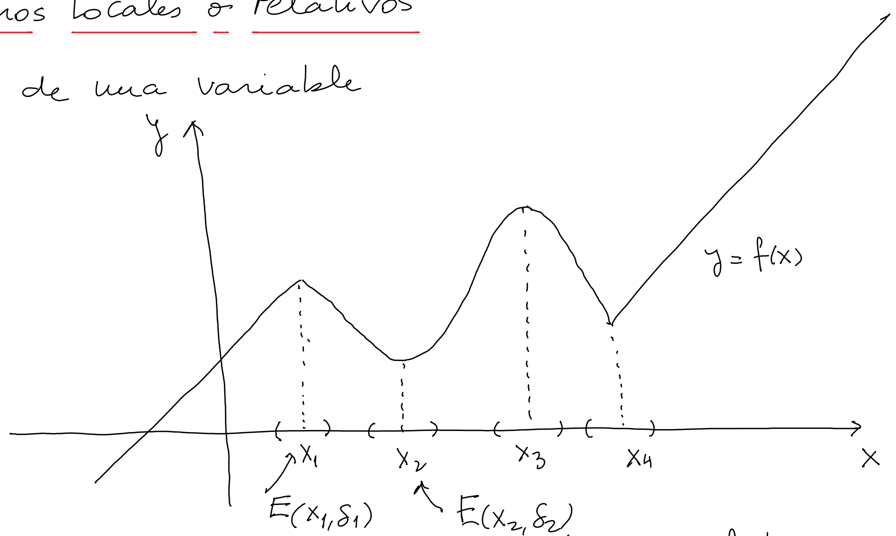
$$= \frac{1}{2}\left(6x^2+\underbrace{x(y-2)+x(y-2)}_{2x(y-2)}\right)$$

$$= 3x^2+x(y-2)$$

$$P_2(x,y) = \overbrace{2x+y}^{P_1(x,y)} + 3x^2 + x(y-2) = \cancel{2x}+y+3x^2+xy-\cancel{2x} = y+3x^2+xy$$

Extremos locales o relativos

funciones de una variable



En los puntos x_1 y x_3 f alcanza máximos relativos, mientras que en x_2 y x_4 alcanza mínimos relativos.

En x_1 se produce un máximo relativo porque existe un entorno $E(x_1,\delta_1)$ en el cual se cumple que

$$f(x) \leq f(x_1) \quad \forall\, x \in E(x_1,\delta_1)$$

Lo mismo ocurre en x_3.

En x_2 se produce un mínimo relativo porque hay un entorno $E(x_2,\delta_2)$ tal que

$$f(x) \geq f(x_2) \quad \forall x \in E(x_2,\delta_2)$$

y lo mismo ocurre en x_4.

En general, decimos que f alcanza un máximo local o relativo en un punto x_M si existe un entorno $E(x_M,\delta) \subset \mathrm{dom}(f)$ en el cual se cumple que f restringida a $E(x_M,\delta)$ tiene un máximo en x_M: $f(x) \leq f(x_M) \quad \forall x \in E(x_M,\delta)$

Análogamente, f alcanza un mínimo relativo en x_m si existe

Análogamente, f alcanza un mínimo relativo en x_m si existe un entorno $E(x_m, \delta)$ tal que f restringida a $E(x_m,\delta)$ tiene un mínimo en x_m: $f(x) \geq f(x_m) \quad \forall x \in E(x_m, \delta)$

Note que en los puntos x_1, x_2, x_3 y x_4, en los cuales f alcanza máximos o mínimos locales, f no tiene derivada (punto x_1, x_4) o la derivada existe y es nula (puntos x_2 y x_3)

Se dice que f alcanza un extremo local o relativo en x_0 si en ese punto se produce un máximo o mínimo relativo.

Extremos locales o relativos de campos escalares

Sea $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ abierto.

f alcanza un máximo relativo en $x_0 \in D$ si existe un $E(x_0, \delta)$ tal que

$$f(x) \leq f(x_0) \qquad \forall x \in E(x_0, \delta)$$

es decir, f restringida al entorno $E(x_0,\delta)$ alcanza un máximo en x_0.

f alcanza un mínimo relativo en $x_0 \in D$ si existe un $E(x_0, \delta)$ tal que

$$f(x) \geq f(x_0) \qquad \forall x \in E(x_0, \delta)$$

es decir, f restringida al entorno $E(x_0,\delta)$ alcanza un mínimo en x_0.

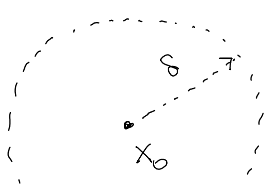
Decimos que f alcanza un extremo relativo en x_0 si allí se produce un máximo o mínimo relativo y $f(x_0)$ es el valor del extremo.

Condición necesaria (pero no suficiente) de extremo relativo

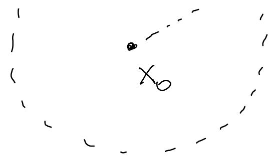
Teorema: Sea $f: D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ abierto. Si f alcanza un extremo relativo en $x_0 \in D$ y es diferenciable en ese punto, entonces

$$\nabla f(x_0) = \vec{0}$$

Demostración: Sea $E(x_0,\delta)$ el entorno en el cual f restringida a ese entorno alcanza un máximo o mínimo en x_0.



$f(x) \geq f(x_0) \quad \forall x \in E(x_0, \delta)$ si se produce un mínimo en x_0, o

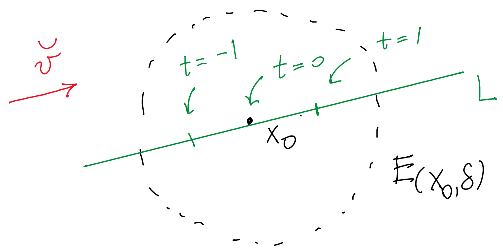


un mínimo en x_0, o
$f(x) \leq f(x_0)$ $\forall x \in E(x_0, \delta)$ si se produce un máximo en x_0

Sea $\breve{v}$ un versor. Consideremos la recta

$$L: X = X_0 + t\breve{v} \qquad t \in \mathbb{R}$$

y llamemos $g(t) = f(X_0 + t\breve{v})$. Notamos que $g(0) = f(X_0)$. Entonces



g alcanza un máximo relativo en $t=0$ si f produce un máximo relativo en X_0, y g tiene un mínimo relativo en $t=0$ si f alcanza un mínimo relativo en X_0.

Como f es diferenciable en X_0 y $\vec{\alpha}(t) = X_0 + t\breve{v}$ es diferenciable en $t=0$, $g(t) = f \circ \vec{\alpha}(t)$ es derivable en $t=0$. Luego

$$g'(0) = Df(\vec{\alpha}(0)) \cdot \vec{\alpha}'(0) = Df(X_0) \cdot \breve{v} = \nabla f(X_0) \cdot \breve{v}$$

Como g alcanza un extremo relativo en $t=0$, $g'(0)=0$. Luego

$\nabla f(X_0) \cdot \breve{v} = 0 \quad \forall \breve{v} \Rightarrow \nabla f(X_0) \perp \breve{v}$ para todo versor $\breve{v}$.

Luego, $\nabla f(X_0) = \vec{0}$, pues $\vec{0}$ es el único vector ortogonal a todo posible versor.

Interpretación geométrica: Si $f(x,y)$ alcanza un extremo relativo en (x_0, y_0) y es diferenciable en ese punto,

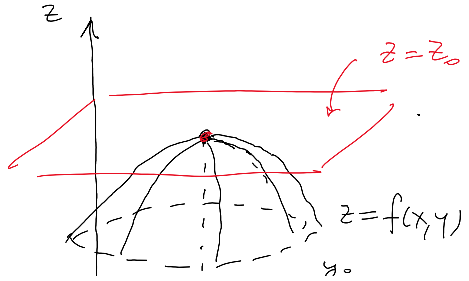
$$\nabla f(x_0, y_0) = (f'_x(x_0, y_0), f'_y(x_0, y_0)) = (0,0).$$

Entonces el plano tangente al gráfico de f en $(x_0, y_0, z_0) = (x_0, y_0, f(x_0,y_0))$ es

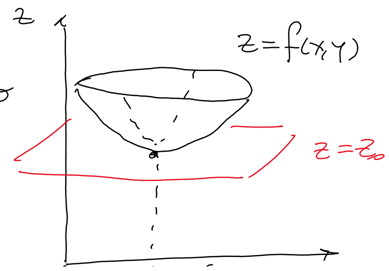
$$z = z_0 + \underbrace{\cancel{f'_x(x_0,y_0)}}_{0}(x-x_0) + \underbrace{\cancel{f'_y(x_0,y_0)}}_{0}(y-y_0)$$

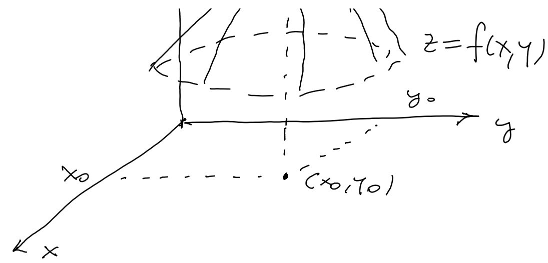
$z = z_0$ (plano paralelo al plano xy)

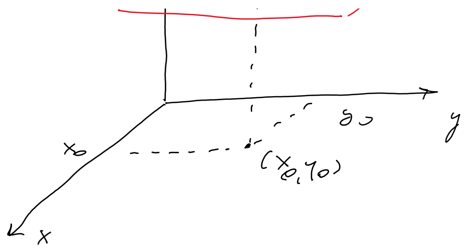
Máximo relativo



mínimo relativo







Puntos críticos o estacionarios

$f:D\to\mathbb{R}$, $D\subset\mathbb{R}^n$, tiene un punto crítico o estacionario en $X_0\in D$ si $X_0\in \text{int}(D)$ y pasa una de las siguientes:

a) f no es diferenciable en x_0

b) f es diferenciable en x_0 y $\nabla f(x_0)=\vec{0}$.

Con esta definición tenemos que los puntos en los cuales se alcanza un extremo relativo son necesariamente puntos críticos:

- f alcanza extremo relativo en $X_0 \Rightarrow X_0$ punto crítico de f

- $\left.\begin{array}{l} f \text{ alcanza extremo relativo en } x_0 \\ \qquad + \\ f \text{ diferenciable en } x_0 \end{array}\right\} \Rightarrow \nabla f(x_0)=\vec{0}$

Luego, si nos interesa hallar extremos relativos, debemos comenzar por hallar los puntos críticos.

No todo punto crítico es necesariamente productor de un extremo relativo.

Punto silla: punto crítico en el cual no se produce extremo relativo

Clasificación de puntos críticos:
- máximo relativo
- mínimo relativo
- punto silla

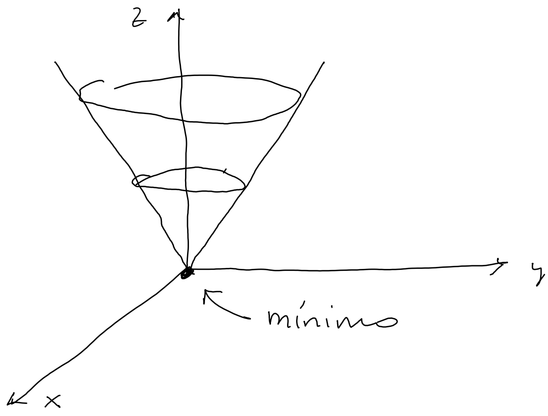
Ejemplos:

1) $f(x,y)=\sqrt{x^2+y^2}$ tiene un punto crítico en $(0,0)$ pues no

Ejemplos:

1) $f(x,y)=\sqrt{x^2+y^2}$ tiene un punto crítico en $(0,0)$ pues no es diferenciable allí. Tal punto produce un mínimo relativo porque

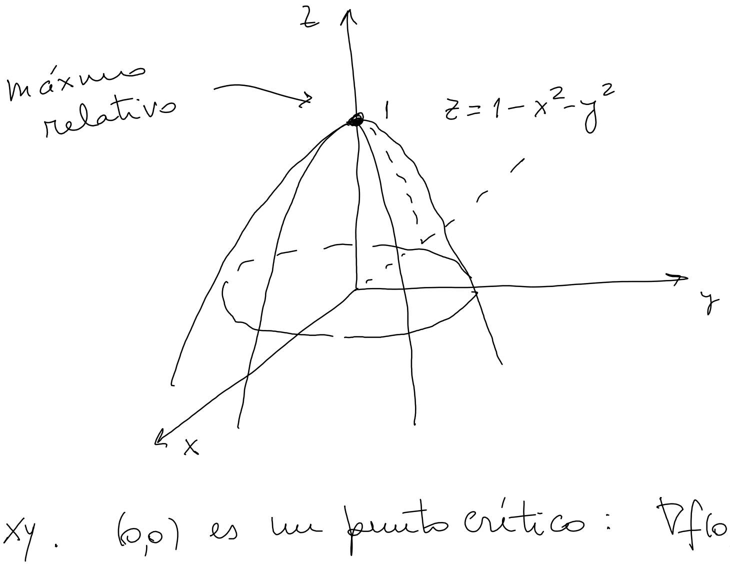
$$f(0,0)=0\leqslant f(x,y)=\sqrt{x^2+y^2}\quad \forall (x,y)\in\mathbb{R}^2$$



2) $f(x,y)=1-x^2-y^2$ tiene un punto crítico en $(0,0)$ porque

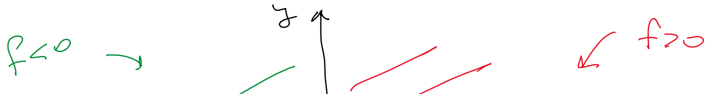
$$\nabla f(0,0)=(-2x,-2y)\Big|_{(0,0)}=(0,0)$$

Como $f(0,0)=1$ y $f(x,y)=1-x^2-y^2\leqslant 1\ \forall (x,y)\in\mathbb{R}^2$,
$(0,0)$ produce un máximo relativo.

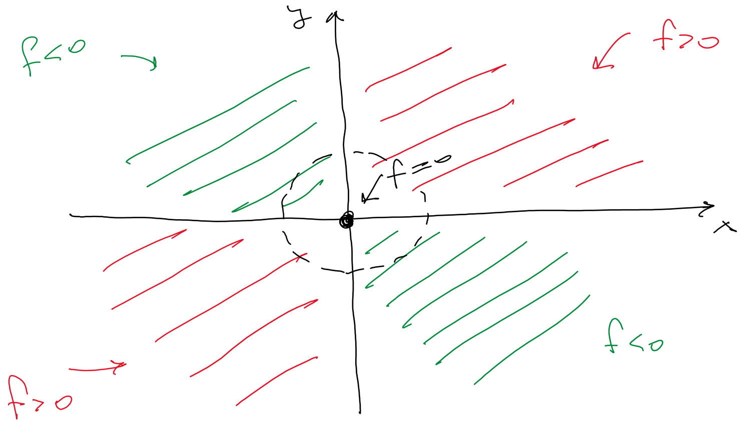


3) $f(x,y)=xy$. $(0,0)$ es un punto crítico: $\nabla f(0,0)=(y,x)\Big|_{(0,0)}=(0,0)$

- $f(0,0)=0$
- $f(x,y)>0 \iff xy>0 \iff (x>0\wedge y>0)\vee(x<0\wedge y<0)$
- $f(x,y)<0 \iff xy<0 \iff (x>0\wedge y<0)\vee(x<0\wedge y>0)$



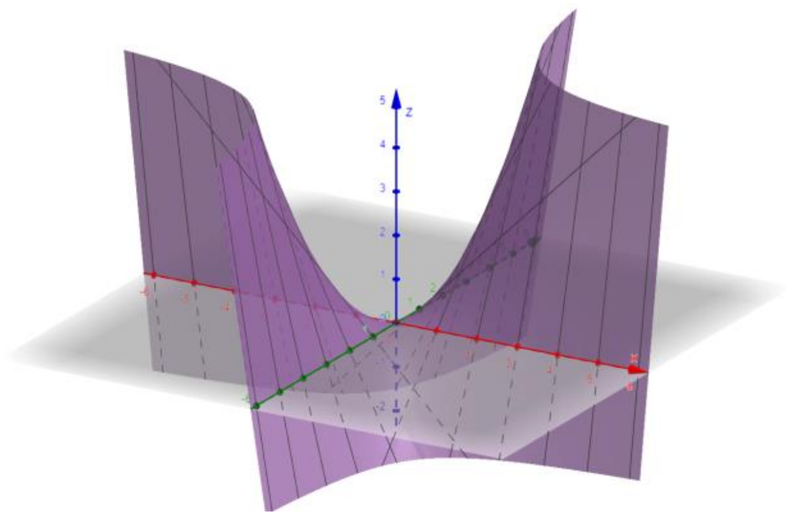
En $(0,0)$ no



En $(0,0)$ no se produce ni un máximo ni un mínimo pues en cada posible $E((0,0),\delta)$ hay puntos con $f>0$ y puntos con $f<0$

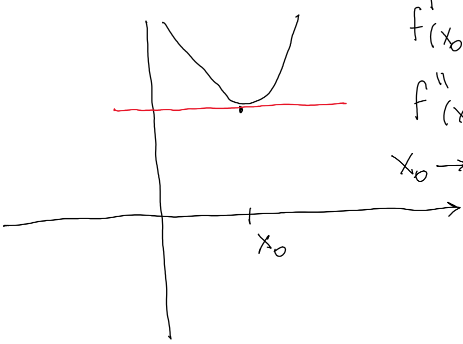
Entonces $(0,0)$ es punto silla (no es extremo)

El gráfico de f tiene forma de silla de montar



Criterio de las segundas derivadas

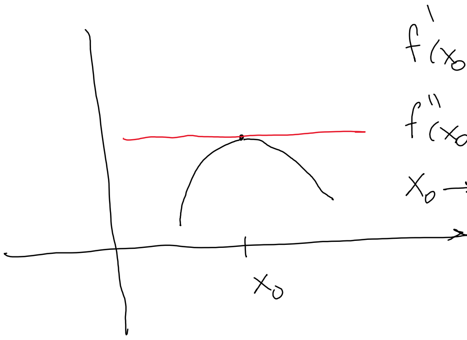
Caso funciones de una variable



$f'(x_0)=0$

$f''(x_0)>0$

$x_0 \to$ mínimo rel.



$f'(x_0)=0$

$f''(x_0)<0$

$x_0 \to$ máximo

Criterio de las 2das derivadas funciones dos variables

Suponemos que $f(x,y)$ es $\mathcal{C}^2$ y $\nabla f(x_0,y_0)=(0,0)$

Suponemos que $f(x,y)$ es $\mathcal{C}^2$ y $\nabla f(x_0,y_0) = (0,0)$

$$Hf(x_0,y_0) = \begin{pmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{pmatrix}$$ es la matriz Hessiana

Llamamos $$\Delta = \det\left(Hf(x_0,y_0)\right) = \begin{vmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{vmatrix}$$

1) Si $f''_{xx}(x_0,y_0) > 0 \ \wedge \ \Delta > 0 \Rightarrow (x_0,y_0)$ produce mínimo relativo

2) Si $f''_{xx}(x_0,y_0) < 0 \ \wedge \ \Delta > 0 \Rightarrow (x_0,y_0)$ produce máximo relativo

3) Si $\Delta < 0 \Rightarrow (x_0,y_0)$ produce punto silla.

4) Si $\Delta = 0$ el criterio no da información

Criterio 2das derivadas funciones 3 variables

$f(x,y,z)$ de clase $\mathcal{C}^2$, $\nabla f(x_0,y_0,z_0) = (0,0,0)$

Llamamos
$$\Delta_2 = \begin{vmatrix} f''_{xx}(x_0,y_0,z_0) & f''_{xy}(x_0,y_0,z_0) \\ f''_{yx}(x_0,y_0,z_0) & f''_{yy}(x_0,y_0,z_0) \end{vmatrix} \quad y \quad \Delta_3 = \det\left(Hf(x_0,y_0,z_0)\right)$$

Entonces:

1) Si $f''_{xx}(x_0,y_0,z_0) > 0$, $\Delta_2 > 0$ y $\Delta_3 > 0 \Rightarrow (x_0,y_0,z_0)$ produce mínimo relativo

2) Si $f''_{xx}(x_0,y_0,z_0) < 0$, $\Delta_2 > 0$ y $\Delta_3 < 0 \Rightarrow (x_0,y_0,z_0)$ produce máximo relativo

3) Si $\Delta_3 \neq 0$ y no se cumplen 1) o 2) $\Rightarrow (x_0,y_0,z_0)$ produce punto silla

4) Si $\Delta_3 = 0$, el criterio no da información.

Por ejemplo si en (x_0,y_0,z_0):

1) $f''_{xx} = 1$, $\Delta_2 = 2$, $\Delta_3 = \frac{1}{2} \longrightarrow$ mínimo

2) $f''_{xx} = -1$, $\Delta_2 = 1$, $\Delta_3 = -2 \longrightarrow$ máximo

2) $f''_{xx} = -1$, $\Delta_2 = 1$, $\Delta_3 = -2 \longrightarrow$ máximo

3) $f''_{xx} = -1$, $\Delta_2 = -1$, $\Delta_3 = 2 \longrightarrow$ punto silla

4) $f''_{xx} = 1$, $\Delta_2 = 0$, $\Delta_3 = 1 \longrightarrow$ punto silla

5) $f''_{xx} = -1$, $\Delta_2 = 2$, $\Delta_3 = 0 \longrightarrow$ no se puede concluir