Cálculo en dos o más varias variables

Los temas que veremos en esta primera parte se aplican en diferentes áreas de la informática como la computación gráfica, análisis de imágenes, e Inteligencia artificial (Minería de datos y Aprendizaje Profundo)

Vamos a estudiar conceptos que se vieron en una variable, pero extendiéndolos ahora a más de una (ya que en general, al estudiar fenómenos del mundo real es usual que una cantidad dependa de más de una variable).

1 Funciones de varias variables - Límites y Continuidad

Definición 1.1. *Una función real f de n variables es una regla que asigna a cada n-ulpa de números reales $(x_1, x_2, \ldots, x_n)$ un único número real $f(x_1, x_2, \ldots, x_n)$.*

Esto se puede escribir de la siguiente forma:

$$f : D \subset \mathbb{R}^n \to \mathbb{R} \tag{1}$$

Donde D es un subconjunto de $\mathbb{R}^n$ en donde está definida la función f, a este conjunto D se lo llama ***dominio*** de f, ***Dom(f)***.

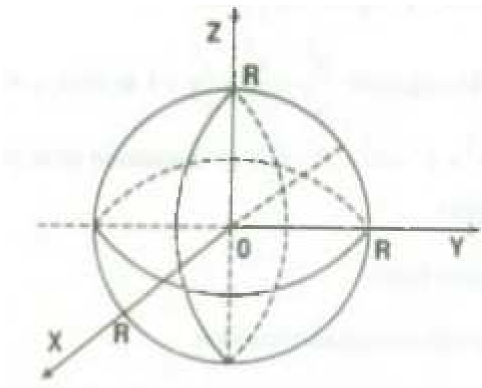
Por otro lado, la ***imagen*** o ***rango*** (de D por f) ***Im(f)***,es un subconjunto de $\mathbb{R}$ formado por todos los valores que toma f en $(x_1, x_2, \ldots, x_n)$, es decir, por $f(x_1, x_2, \ldots, x_n)$.

Ejemplo 1.2. *Sea* $f : D \subset \mathbb{R}^3 \to \mathbb{R}$ *dada por* $f(x, y, z) = \sqrt{1 - x^2 - y^2 - z^2}$*, analizar su dominio e imagen.*

El $Dom(f)$ *está dado por todos los puntos de* $\mathbb{R}^3$ *en donde está definida la función, por lo cual:*

$$
\begin{aligned}
D &= \left\{(x, y, z) \in \mathbb{R}^3 : 1 - x^2 - y^2 - z^2 \geq 0\right\} \\
&= \left\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\right\}
\end{aligned}
$$

El dominio no es más que una bola de radio 1, centrado en $(0, 0, 0)$*, es decir la esfera de radio 1 y su interior.*



Por otro lado, para todo valor de $(x, y) \in D$*, se tiene que* $f(x, y, z) \in [0, 1]$*, por lo tanto* $Im(f) = [0, 1]$.

1.1 Funciones de dos variables

Definición 1.3. *Una función real f de dos variables es una regla que asigna a cada par ordenado de números reales (x, y) un único número real $f(x, y)$.*

En este caso el dominio de f ($Dom(f)$), que es el subconjunto de $\mathbb{R}^2$ en el cual está definida la función, se representa como una región del plano, pues está dado por todos los puntos del plano para los cuales $f(x, y)$ es un número real bien definido.

Por otra parte, la imagen de f es el subconjunto de $\mathbb{R}$ formado por los valores que toma la función f en (x, y).

Si $(x, y) \in D$, se suele escribir $z = f(x, y)$ donde queda explícitamente definido que z es el valor que toma la función f al evaluarla para el par ordenado (x, y).

Ejemplos 1.4.

- *La función nula $f(x, y) = 0$ está definida para todo $(x, y) \in \mathbb{R}^2$, por tanto $D = \mathbb{R}^2$ y su imagen es el conjunto $Im(f) = \{0\}$.*

- *La función constante $f(x, y) = c$ (siendo c una constante fija) tiene como dominio todo $\mathbb{R}^2$ y su imagen es $Im(f) = \{c\}$.*

- *La función $f(x, y) = ax + by + c$, que se denomina función lineal, tiene como dominio todo $\mathbb{R}^2$ y su imagen es $Im(f) = \mathbb{R}$.*

- *Dada la función $f(x, y) = \frac{x}{x^2+y^2}$, determinemos su dominio e imagen:*

  Observamos que f está dada por un cociente, por lo tanto f está bien definida siempre que el denominador sea no nulo. Esto es, $x^2 + y^2 \neq 0$, lo que implica que x e y no pueden ser simultáneamente nulos.

  Por lo tanto $Dom(f) = D = \big\{(x, y) \in \mathbb{R}^2 : (x, y) \neq (0, 0)\big\} = \mathbb{R}^2 - \{(0, 0)\}$.

  Por otro lado dado que z puede tomar cualquier valor real entonces $Im(f) = \mathbb{R}$.

Representación gráfica de funciones

Una forma de visualizar el comportamiento de una función de dos variables es mediante la representación de su gráfica.

Para una función $g(x)$ de una variable recordemos que su gráfica era una curva C en el plano con ecuación $y = g(x)$. La gráfica de una función f de dos variables es una superficie S en el espacio, con ecuación $z = f(x, y)$.

Definición 1.5. *Se llama gráfica de una función f de dos variables al conjunto de todos los puntos del espacio con coordenadas (x, y, z) tales que $(x, y) \in Dom(f)$ y $z = f(x, y)$.*

Si llamamos S a la superficie que representa la gráfica de f, entonces:

$$S = \{(x, y, z) : (x, y) \in Dom(f); z = f(x, y)\} = \{(x, y, f(x, y)) : (x, y) \in Dom(f)\}$$

Ejemplos 1.6.

- *La función nula tiene como gráfica la superficie $z = 0$.*

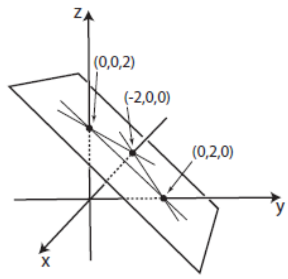
- *La función constante $f(x, y) = c$ se representa gráficamente como el plano (horizontal) de ecuación $z = c$.*

- *La gráfica de la función $f(x, y) = x - y + 2$ es el plano $z = x - y + 2$.*

  Como no hay ninguna condición particular para el dominio de f, consideraremos el dominio natural, $Dom(f) = \mathbb{R}^2$

  Para obtener la imagen observamos que los valores que toma f son $z = x - y + 2$, con $(x, y) \in \mathbb{R}^2$, y como vemos que z puede adoptar cualquier valor real, $Im(f) = \mathbb{R}^2$

  Para trazar la gráfica de f, escribimos $z = f(x, y) = x - y + 2$, que corresponde a la ecuación de un plano.



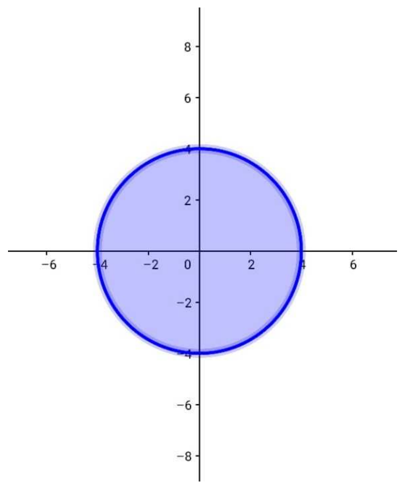
- *Dada la función* $f(x,y) = \sqrt{16 - x^2 - y^2}$, *hallar el dominio, la imagen y su representación gráfica.*

  El $Dom(f)$ *está dado por todos los puntos de* $\mathbb{R}^2$ *en donde está definida la función, por lo cual:*

  $$D \;=\; \left\{(x,y) \in \mathbb{R}^2 : 16 - x^2 - y^2 \geq 0\right\} = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 16\right\}$$

  Corresponde al disco radio 4, centrado en $(0,0)$, *es decir la circunferencia de radio 4 y su interior.*

  Para los puntos fuera de la circunferencia la función no está definida.



  Por otro lado, para todo valor de $(x,y) \in D$, *se tiene que* $f(x,y,z) \in [0,4]$, *pues para los puntos de la frontera del disco se cumple que* $16 - x^2 - y^2 = 0$ *y para el origen* $f(0,0) = \sqrt{16} = 4$, *por lo tanto* $Im(f) = [0,4]$.

$$
\begin{aligned}
Graf(f) &= \{(x,y,z) : (x,y) \in Dom(f); z = f\,(x,y)\} = \left\{(x,y,z) : z = \sqrt{16 - x^2 - y^2}; (x,y) \in D\right\} \\
&= \left\{(x,y,z) : z^2 = 16 - x^2 - y^2; (x,y) \in D\right\} = \left\{(x,y,z) : x^2 + y^2 + z^2 = 16; (x,y) \in D\right\}
\end{aligned}
$$

La gráfica corresponde a la (semi)esfera (superior) con centro en el origen y radio 4.

La superficie completa es la esfera pero no puede ser la gráfica de una función (recordar la definición de función, para cada elemento del dominio le corresponde un único elemento en el codominio). Identificamos entonces que la gráfica de la función $f(x,y)$ con la mitad superior de la superficie de la esfera con centro en el origen y radio 4 (con $z \geq 0$).

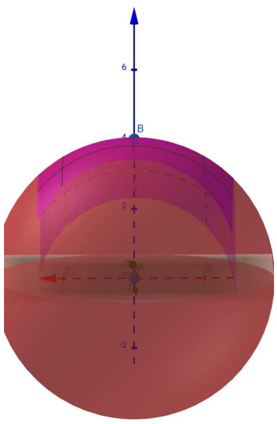


Figure 1: Esfera de radio 4 con centro en (0,0,0)

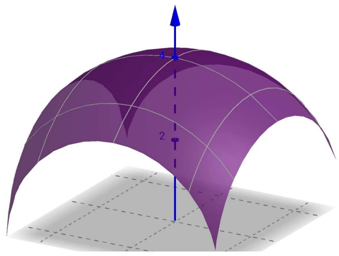


Figure 2: Semiesfera radio 4 con centro en (0,0,0)

Curvas de nivel

Se llama *curva de nivel* k de una función f de dos variables al conjunto de todos los puntos del dominio de la función con coordenadas (x, y) tales que $f(x, y) = k$, donde k es una constante que pertenece a la imagen de f. Llamando C_k a la curva de nivel k, entonces para cada $k \in Im(f)$.

$$C_k = \{(x, y) : (x, y) \in Dom(f); f(x, y) = k\}$$

La manera de representar a la función es mediante su *mapa de niveles o mapa de contornos*, que se obtiene dibujando unas cuantas curvas de nivel, para distintos valores de k. Es común tomar valores de k equiespaciados.

Por construcción, para los pares del dominio que forman una dada curva de nivel, la función f toma el mismo valor. Luego, la curva de nivel k muestra todos los pares del dominio donde la gráfica de f tiene nivel o *altura* k.
A partir de las curvas de nivel rotuladas con su nivel o altura de función, se puede inferir la gráfica de la función, elevando mentalmente cada curva de nivel hasta la altura apropiada.
Si se hiciera este procedimiento para todas las curvas de nivel C_k con $k \in Im(f)$, juntas conformarán la gráfica de f.

La curva de nivel C_k de una función $f(x, y)$ es precisamente la proyección en el plano xy de la traza horizontal $z = k$ de la superficie que es gráfica de f.
O sea que, si se dibujan curvas de nivel de una función y se visualizan como si se elevaran hasta el nivel que indica k, es posible trazar mentalmente una gráfica aproximada.

No haremos más hincapié en el estudio de las curvas de nivel, podemos identificar superficies usando GeoGebra y otros graficadores.

1.2 Límites y Continuidad

El concepto de límite es una herramienta básica y útil para el análisis de funciones.

Intuitivamente, la idea de *límite de f cuando x tiende (se acerca) a* x_0 *es igual a* L significa que a medida que x se acerca a x_0 los valores que va tomando f se acercan cada vez más al valor de L.

Al ser f una función de una variable el dominio está incluido en la recta real, (es unidimensional), sólo hay dos direcciones o caminos posibles para llegar al x_0: desde la izquierda o desde la derecha.
Si el límite por la izquierda es distinto del límite por la derecha, entonces el límite de la función cuando x se acerca a x_0 no existe, mientras que si ambos límites laterales existen y coinciden entonces la función tiene ese lḿite.

Para funciones de varias variables el concepto de límite es similar al visto para una variable, aunque el cálculo es un poco más complejo.

Se dice que una función de dos variables $f(x, y)$ tiene límite L (un número real fijo) cuando (x, y) tiende a (x_0, y_0), si para todos los puntos (x, y) cercanos al punto (x_0, y_0), los valores $f(x, y)$ son arbitrariamente próximos al número L.

La definición es similar a la del límite funcional para una variable. Sin embargo, para el caso de una función de dos variables (definida sobre un dominio bidimensional $D \subset R^2$), el punto (x, y) podrá acercarse al punto (x_0, y_0) desde muchas direcciones. De hecho, hay infinitas maneras de acercarse a un punto en el plano.

Definición 1.7. *Se dice que una función $f(x,y)$ tiene límite L si cuando (x,y) tiende a (x_0,y_0) la función $f(x,y)$ se acerca al valor L y se escribe* $\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$

Formalmente el límite existe si dado un número ϵ existe un número $\delta > 0$, tal que $\forall(x,y) \in Dom(f)$, entonces

$$0 < \sqrt{(x-x_0)^2+(y-y_0)^2} \leq \delta \to |f(x,y)-L| < \epsilon$$

La diferencia $|f(x,y)-L|$ es la distancia entre los números $f(x,y)$ y L (en la recta); mientras que $\sqrt{(x-x_0)^2+(y-y_0)^2}$ es la distancia entre el punto $(x;y)$ en el dominio de la función y el punto (x_0,y_0) (en el plano).

Entonces, la definición dice que la distancia entre $f(x,y)$ y L es arbitrariamente pequeña siempre que la distancia entre (x,y) y (x_0,y_0) sea suficientemente pequeña (aunque no nula). El punto (x_0,y_0) puede no pertenecer al $Dom(f)$, el único requisito es que los puntos varíen en el $Dom(f)$.

Ejemplo 1.8. *Calculemos un límite por definición :*

$$\lim_{(x,y)\to(0,0)} \frac{x.y^2}{x^2+y^2} \qquad (*)$$

Sea $\epsilon > 0$, queremos encontrar un valor $\delta > 0$ tal que:

si $\underbrace{0 < \sqrt{(x-0)^2+(y-0)^2} < \delta}_{\sqrt{x^2+y^2}<\delta}$ *entonces* $\left|\frac{x.y^2}{x^2+y^2} - 0\right| < \epsilon$

Tenemos que $\left|\frac{x.y^2}{x^2+y^2}\right| = \left|x\frac{y^2}{x^2+y^2}\right| = |x| \underbrace{\frac{y^2}{x^2+y^2}}_{\leq 1} \leq |x| = \sqrt{x^2} \leq \sqrt{x^2+y^2} < \delta = \epsilon$

Luego, para cualquier ϵ existe un $\delta = \epsilon$ tal que vale (*) como queríamos

La definición de límite de una función sólo hace referencia a la distancia entre los puntos (x,y) y (x_0,y_0), nada dice sobre la dirección de acercamiento. Por tanto, si existe el límite, $f(x,y)$ debe acercarse al mismo valor L independientemente de como el punto (x,y) se acerque a (x_0,y_0). Pero resulta imposible, obviamente, analizar todos los caminos que llegan a (x_0,y_0)

para ver a qué valor tiende f por cada uno de ellos.

Para determinar si una función $f(x, y)$ tiene o no límite uno elige dos o tres caminos que nos lleven hacia el punto (x_0, y_0), si resulta que los valores obtenidos para ese límite son distintos según el camino elegido entonces ese límite **NO** existe, y por tanto se tiene un criterio para determinar la no existencia del límite.

Sin embargo, si se prueba por varios caminos y se obtiene el mismo valor de L esto no alcanza para asegurar que el límite exista y sea ese valor pero nos permite *suponer* que el límite existe y toma ese valor L para luego demostrarlo efectivamente usando la definición o algunas propiedades.

A menudo demostrar la existencia del límite resulta a partir de la definición resulta difícil, sin embargo existen diferentes herramientas alternativas como por ejemplo el teorema del encaje o *criterio del "'sandwich"'* que establece:

Teorema 1.9. *Si existen funciones $g(x, y)$ y $h(x, y)$ tales que*

$$g(x, y) \leq f(x, y) \leq h(x, y)$$

$\forall (x, y) \neq (x_0, y_0)$ en un disco con centro en (x_0, y_0), y si

$$\lim_{(x,y)\to(x_0,y_0)} g(x, y) = \lim_{(x,y)\to(x_0,y_0)} h(x, y) = L$$

entonces

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

No demostraremos el teorema ni las próximas propiedades *(si siente curiosidad puede consultar en clase o ver la demostración en los libros que marcamos como bibliografía recomendada)*

Propiedades 1.10. *Sea $c \in \mathbb{R}$ una constante, y sean $f(x, y)$ y $g(x, y)$ dos funciones rales de dos variables tales que existen los siguientes límites*

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L \qquad y \lim_{(x,y)\to(x_0,y_0)} g(x, y) = M$$

Entonces se cumplen las siguientes propiedades:

- $\lim_{(x,y)\to(x_0,y_0)} (f(x,y)+g(x,y)) = L+M$
- $\lim_{(x,y)\to(x_0,y_0)} cf(x,y) = c.L$
- $\lim_{(x,y)\to(x_0,y_0)} (f(x,y).g(x,y)) = L.M$
- $\lim_{(x,y)\to(x_0,y_0)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{L}{M}$ *si* $M \neq 0$

Además, si $M=0$ *y* $L \neq 0$ *entonces* $\lim_{(x,y)\to(x_0,y_0)} \dfrac{f(x,y)}{g(x,y)}$ *no existe.*

Ejemplos 1.11. *1. Dada la función* $f(x,y) = x^2y+x^2+sen(y^2)$*, calcular* $\lim_{(x,y)\to(1,0)} f(x,y)$*.*

Se trata de una función que está definida en $(1,0)$*, el punto pertenece al dominio, aplicando las propiedades y por evaluación directa:*

$$\lim_{(x,y)\to(1,0)} x^2y+x^2+sen(y^2) = 1^20+1^2+sen(0^2) = 0+1+sen(0) = 1$$

2. Dada la función $f(x,y) = \frac{5x^2y}{x^2+y^2}$*, calcular* $\lim_{(x,y)\to(2,-1)} f(x,y)$*.*

Se trata de una funcióon racional que está bien definida en $(2,-1)$*. Por lo tanto podemos calcular el límite aplicando la regla del cociente, y por evaluación directa de los polinomios del numerador y denominador:*

$$\lim_{(x,y)\to(2,-1)} \frac{5x^2y}{x^2+y^2} = \frac{52^2(-1)}{2^2+(-1)^2} = -4$$

3. No existe el límite en el $(0,0)$ *de* $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$*, ya que si nos acercamos al origen por el eje* x *la función se acerca a* 1*, pero si nos acercamos al* $(0,0)$ *por el eje* y *la* $f(x,y)$ *se acerca a* -1*. Encontramos dos trayectorias que llegan al origen, pero tales que a lo largo de cada una de ellas* f *toma valores diferentes.*

4. Podemos ver que por iterados el límite $\lim_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2} = 0$

Pero si consideramos llegar al origen a lo largo de una recta de $y = mx$ (que pasa por el punto $(0,0)$), resulta que :

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2} = \lim_{x\to 0} \frac{xmx}{x^2+(mx)^2} = \frac{m}{1+m^2}$$

Lo que nos dará un límite diferente para cada recta con pendiente diferente, entonces no existe el límite.

5. *Se puede probar que la función $f(x,y) = \frac{xy^2}{x^2+y^4}$ tiende a 0 por diferentes trayectorias (los ejes,las rectas que pasan por el $(0,0)$) pero tiende a otros números si nos acercamos al origen por otras curvas. Por ejemplo, si vamos por $y = \sqrt{x}$ el límite será $\frac{1}{2}$.*

6. *Dada la función $f(x,y) = \frac{5x^2y}{x^2+y^2}$, calcular $\lim\limits_{(x,y)\to(0,0)} f(x,y)$.*

Analicemos qué ocurre con la función cuando nos acercamos al punto $(0,0)$ por diferentes caminos:

- ***Probamos Límites iterados**:*

$$\lim_{x\to 0}\left(\lim_{y\to 0} \frac{5x^2y}{x^2+y^2}\right) = \lim_{x\to 0}\left(\frac{5x^2\cdot 0}{x^2+0^2}\right) = \lim_{x\to 0} 0 = 0$$

$$\lim_{y\to 0}\left(\lim_{x\to 0} \frac{5x^2y}{x^2+y^2}\right) = \lim_{y\to 0}\left(\frac{5\cdot 0^2y}{0^2+y^2}\right) = \lim_{y\to 0} 0 = 0$$

- ***Por la recta** $y = mx$*

$$\lim_{x\to 0} \frac{5x^2mx}{x^2+(mx)^2} = \lim_{x\to 0} \frac{5mx^3}{(m+1)x^2} = \lim_{x\to} \frac{5m}{m+1}x = 0$$

- ***Por la curva** $y = x^2$*

$$\lim_{x\to 0} \frac{5x^2x^2}{x^2+(x^2)^2} = \lim_{x\to 0} \frac{5x^4}{x^2(1+x^2)} = \lim_{x\to} \frac{5x^2}{1+x^2} = 0.$$

Como vemos que por distintas trayectorias el límite es el mismo, probaremos con el criterio que es efectivamente ese límite es 0

Teorema del Encaje

$$|\frac{5x^2y}{x^2+y^2}| = 5|y|\frac{x^2}{x^2+y^2} \leq 5|y|$$

de donde,

$$-5|y| \leq \frac{5x^2y}{x^2+y^2} \leq 5|y|$$

Luego, si tomamos $g(x,y) = -5|y|$ *y* $h(x,y) = 5|y|$ *y dado que ambas tienden a* 0 *cuando* $(x,y) \to (0,0)$ *el teorema del encaje nos asegura que:*

$$\lim_{(x,y)\to(0,0)} \frac{5x^2y}{x^2+y^2} = 0$$

Coordenadas Polares

Las coordenadas polares pueden servirnos para calcular límites.

En efecto, si recordamos la relación entre las coordenadas polares y las coordenadas cartesianas.

$$x = r\cos(\theta) \qquad\qquad y = r\sin(\theta)$$

podemos decir que $(x, y) \to (0, 0)$, por ejemplo, equivale a decir (en coordenadas polares) $r \to 0$ (independientemente del valor de θ)

Ejemplo 1.12. *Sea* $f(x, y) = \frac{xy^2}{x^2+y^2}$, *queremos calcular el* $\lim\limits_{(x,y)\to(0,0)} f(x, y)$

Recordando que en coordenadas polares tenemos que $x = r\cos(\theta) \quad y = r\sin(\theta)$,

entonces podemos escribir: $x^2 + y^2 = r^2\cos^2(\theta) + r^2\sin^2(\theta) = r^2\left\{\cos^2(\theta) + \sin^2(\theta)\right\} = r^2$, *luego,*

$$\frac{xy^2}{x^2+y^2} = \frac{rcos(\theta)\, r^2 sen^2(\theta)}{r^2 cos^2(\theta) + r^2 sen^2(\theta)} = \frac{r^3 cos(\theta)\, sen^2(\theta)}{r^2} = rcos(\theta)\, sen^2(\theta)$$

por lo cual:

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^2} = \lim_{r\to 0} r.cos(\theta)\, sen^2(\theta) = 0$$

independientemente del valor de θ

Sabemos de Matemática 2 que el concepto de función continua está asociado a la idea intuitiva de una función cuya gráfica es una curva que se puede dibujar sin levantar el lápiz del papel, esto es, una curva sin saltos.

Esta idea de continuidad se generaliza a funciones de varias variables, así para funciones de dos variables el concepto de continuidad está basado en la noción intuitiva de una función cuya gráfica es una superficie sin huecos ni rupturas.

Definición 1.13. *Una función real* $f(x, y)$ *es continua en un punto* (x_0, y_0) *si:*

- $\exists f(x_0, y_0)$.
- $\exists \lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$
- *Se verifica que* $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$

Decimos que $f(x, y)$ es continua en una región plana R si es continua $\forall (x_0, y_0) \in R$.

La continuidad de una función en un punto significa, intuitivamente, que si se *cambia* el punto (sus coordenadas) en una pequeña cantidad, entonces el valor de la función *cambia* en una pequeña cantidad.

Usando las propiedades de los límites se puede mostrar que las sumas, productos y cocientes, ascómo la composición, de funciones continuas son continuas en sus dominios.

Por ejemplo, una función polinomial de dos variables es continua en todo R^2, la exponencial, seno o coseno de cualquier polinomio en x e y también son funciones continuas en R^2.

Teorema 1.14. *Si* c *es un número real y* f *y* g *son funciones continuas en* (x_0, y_0), *entonces las funciones siguientes son continuas en* (x_0, y_0)*:*

1. *Múltiplo escalar:* cf
2. *Suma (y diferencia):* $f + g$
3. *Producto:* fg
4. *Cociente:* $\frac{f}{g}$ *, si* $g(x_0, y_0) \neq 0$.

Teorema 1.15. *Si h es continua en (x_0, y_0) y g es continua en $h(x_0, y_0)$, entonces la función compuesta $(g \circ h)(x, y) = g(h(x, y))$ es continua en (x_0, y_0).*

Ejemplo 1.16. *Estudiar la continuidad de las funciones: $f(x, y) = \frac{5yx^2}{x^2+y^2}$ y $g(x, y) = \frac{x^2-y^2}{x^2+y^2}$*

El $Dom(f) = R^2 - (0, 0)$ y observamos que f es continua en todos los puntos de su dominio, puesto que es una función racional.

Ahora nos preguntamos si podemos extender esta función de manera tal de lograr una (nueva) función $F(x, y)$ continua en todo $\mathbb{R}^{\nvDash}$ y que coincida con $f(x, y)$ siempre salvo en el origen. Para ello lo que hace falta es dar un valor apropiado para la función F en el origen: el valor apropiado será el límite de f en el origen, que ya vimos que existe.

$$F(x, y) = \begin{Bmatrix} \frac{5yx^2}{x^2+y^2} & si & (x, y) \neq (0, 0) \\ 0 & si & (x, y) = (0, 0) \end{Bmatrix}$$

Dado que g es una función racional, es continua en su dominio que es el conjunto $Dom(g) = R^2 - (0, 0)$. La función g es discontinua *en $(0, 0)$ porque* ***no*** *está* definida *en ese punto, y nos preguntamos si podemos extenderla con continuidad a una nueva función G como en el caso anterior.*
En este caso, como vimos en un ejemplo anterior, no existe el límite en el origen. Por lo tanto no se puede definir una función G de manera de extender la continuidad de g a todo el plano.

2 Diferenciabilidad

Intuitivamente podemos ver que la gráfica de una función continua no puede estar “quebrada”, pero ¿qué pasa si la función es derivable?, ¿qué características adicionales tiene su gráfica? Sabemos de Matemática 2 que la derivada se relaciona con la pendiente de la recta tangente a la gráfica de la función.

Nos preguntamos ahora ¿cómo se extienden estos conceptos a funciones de más variables? ¿cómo podemos, por ejemplo, analizar el “cambio” de una función de dos variables cuando éstas cambian? En principio, podemos mirar como afecta a la función un “cambio parcial”, moviendo las variables de a una. Introducimos entonces el concepto de derivación parcial.

Observación 2.1. *Vamos a trabajar con funciones de dos variables para ayudarnos con la visualización geométrica. Todo es aplicable y se puede extender a funciones de tres o más variables pero no se puede hacer un análisis geométrico de la situación (si no como algo más complejo y abstracto)*

2.1 Derivadas Parciales

Consideremos una función $f(x, y)$ definida sobre $D \subset \mathbb{R}^2$.
Si fijamos una variable, por ejemplo a $y = b$, y permitimos que la otra varíe nuestra función f se convierte en una función de una sola variable y podemos trazar en $\mathbb{R}^3$ la curva $C1 : z = f(x, b)$ que corresponde a la gráfica de la función $F_1(x) = f(x, b)$ de la variable x
(De igual forma si fijamos $x = a$ y la función f queda $F_2(x) = f(a, y)$ siendo gráfica de la curva $C2 : z = f(a, y)$).

Si $F_1(x) = f(x, b)$ tiene derivada en a, entonces a esta derivada se la llama derivada parcial de f con respecto a x en (a, b), y se denota como $\frac{\partial f(a,b)}{\partial x} = F_1'(a)$.
Recordando la definición de derivada de una función de una variable a través del límite del cociente incremental, tenemos:

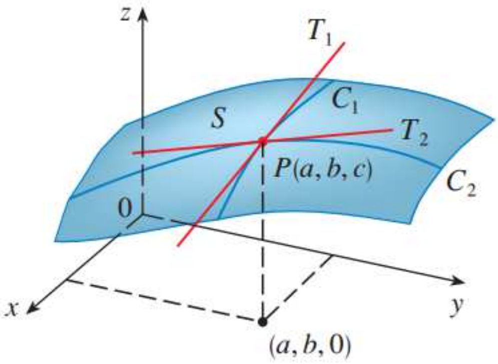


Figure 3: Derivadas Parciales

$$
\begin{aligned}
\frac{\partial f(a,b)}{\partial x} &= F_1'(a) = \lim_{h\to 0}\frac{F_1(a+h)-F_1(a)}{h} \\
&= \lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}
\end{aligned}
$$

Observamos que este cociente incremental se construye evaluando la función en dos puntos próximos, con el mismo valor de $y = y_0$ y dos valores de x próximos en torno a x_0.

Análogamente con $F_2'(b) = \frac{\partial f(a,b)}{\partial y}$

Definición 2.2. *Si $f(x,y)$ es una función de dos variables, sus derivadas parciales respecto de x y de y son las funciones definidas por:*

$$
\begin{aligned}
\frac{\partial f(x,y)}{\partial x} &= \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h} \\
\frac{\partial f(x,y)}{\partial y} &= \lim_{k\to 0}\frac{f(x,y+k)-f(x,y)}{k}
\end{aligned}
$$

si los límites existen.

También usaremos la siguiente notación para las derivadas parciales:

$$f_x(x,y) = \frac{\partial f(x,y)}{\partial x}, \quad f_y(x,y) = \frac{\partial f(x,y)}{\partial y}.$$

Observación 2.3. *En analogía con lo que sucede con funciones de una variable, $\frac{\partial f(x,y)}{\partial x}$ representa la razón de cambio instantánea de f con respecto a x cuando y se mantiene fija, es decir cuando el punto (x, y) se mueve en la dirección del vector $e_1 = (1, 0)$, y equivalentemente $\frac{\partial f(x,y)}{\partial y}$ corresponde a la razón de cambio instantánea cuando el punto se mueve en la dirección de $e_2 = (0, 1)$, esto nos permite decir que las derivadas parciales de f respecto de x e y son las derivadas en las direcciones de los vectores de la base canónica de $\mathbb{R}^2$.*

Ejemplo 2.4. *Calculemos las derivadas parciales de $f(x, y) = x^2 y$ en el punto$(-1, 2)$*

- $\frac{\partial f(-1,2)}{\partial x} = lim_{h\to 0}\frac{f(-1+h,2)-f(-1,2)}{h} = lim_{h\to 0}\frac{(-1+h)^2 2-(-1)^2 2}{h} = lim_{h\to 0}\frac{(1-2h+h^2)2-2}{h} =$

  $= lim_{h\to 0}\frac{2+2h^2-4h-2}{h} = lim_{h\to 0}\frac{h(2h-4)}{h} = lim_{h\to 0}(2h-4) = -4$

- $\frac{\partial f(-1,2)}{\partial y} = lim_{k\to 0}\frac{f(-1,2+k)-f(-1,2)}{k} = lim_{k\to 0}\frac{(-1)^2(2+k)-(-1)^2 2}{k} = lim_{k\to 0}\frac{2+k-2}{k} =$

  $= lim_{k\to 0}\frac{k}{k} = lim_{k\to 0}1 = 1$

A los fines prácticos para calcular las derivadas parciales es posible aplicar las reglas de derivación válidas para funciones de una variable (manteniendo a la otra fija, como si fuera una constante). Entonces, si las derivadas parciales son funciones continuas en $D \subset \mathbb{R}^2$, para obtener $f_x(x_0, y_0)$ y $f_y(x_0, y_0)$ se puede derivar por regla y luego evaluar las expresiones de f_x y f_y en el punto $(x_0, y_0) \in D$.

Ejemplos 2.5. *Calculemos las siguientes derivadas parciales utilizando reglas de derivación:*

1. $f(x, y) = x^2y$ *en el punto*$(-1, 2)$

   - $\frac{\partial f(x,y)}{\partial x} = \frac{\partial (x^2y)}{\partial x} = 2xy$; $\quad \frac{\partial f(-1,2)}{\partial x} = 2(-1)2 = -4$
   - $\frac{\partial f(x,y)}{\partial y} = \frac{\partial (x^2y)}{\partial y} = x^2$; $\quad \frac{\partial f(-1,2)}{\partial x} = (-1)^2 = 1$

2. $g(x, y) = xe^{xy}$ *en el punto* $(1, 0)$

   - $\frac{\partial g(x,y)}{\partial x} = \frac{\partial (xe^{xy})}{\partial x} = e^{xy} + xe^{xy}y = e^{xy}(1 + xy)$; $\quad \frac{\partial g(1,0)}{\partial x} = e^{1.0}(1 + 1.0) = e^0.1 = 1$
   - $\frac{\partial g(x,y)}{\partial y} = \frac{\partial (xe^{xy})}{\partial y} = xe^{xy}x = x^2e^{xy}$; $\quad \frac{\partial g(1,0)}{\partial y} = 1^2e^{1.0} = 1e^0 = 1.1 = 1$

2.2 Diferenciabilidad para funciones de dos variables

Si tenemos una función de una variable diferenciable (derivable) en un punto podemos asegurar que al acercarnos suficientemente a ese punto de la gráfica, la curva que forma la gráfica no se distingue de la *recta tangente* en dicho punto, y podemos aproximar localmente a la función con una función lineal. (ver Figure 1. cada una de las curvas $C1$ y $C2$ admiten rectas tangentes $T1$ y $T2$ respectivamente, en el punto (a,b)).

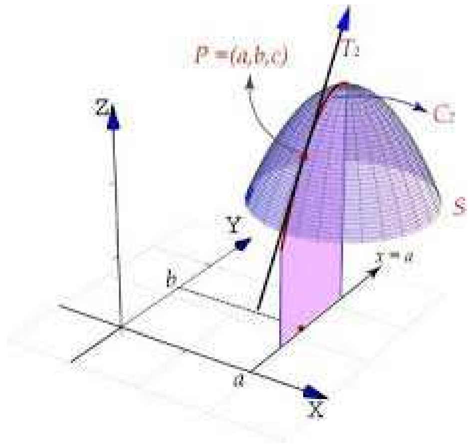
En este contexto nos interesa extender el concepto de diferenciabilidad en un punto para una función de dos variables y poder construir la *aproximación lineal* de tal función.
Dicho en términos geométricos, queremos que al acercarnos suficientemente a un punto de la gráfica de una función (en $\mathbb{R}^3$), la superficie que forma su gráfica no se distinga del plano tangente en dicho punto, y entonces podamos aproximar localmente la función (de dos variables) mediante una función lineal de dos variables (la que corresponde al plano tangente).

Para analizar el tema de la diferenciabilidad para funciones de dos variables supongamos primero que existe el plano tangente a la gráfica S de una función $f(x, y)$ en un punto $P = (a, b, f(a, b))$ y analicemos como debe ser la ecuación de dicho plano.

Para ello recordamos que en una variable, la recta tangente en un punto (x_0, y_0) está dada por $y = f'(x_0)(x - x_0) + f(x_0)$.

Generalizando esto a dos dimensiones, el plano tangente [1] a la superficie S en un punto P, deberá contener a las rectas tangentes en P a cada una de las curvas que están en S y pasan por P.

Las trazas C_1 y C_2 para $y = b$ y $x = a$ en S, respectivamente, son curvas que están en S y pasan por P; además la recta que es tangente a cada una de esas curvas en P tiene como pendiente la derivada parcial correspondiente de f.

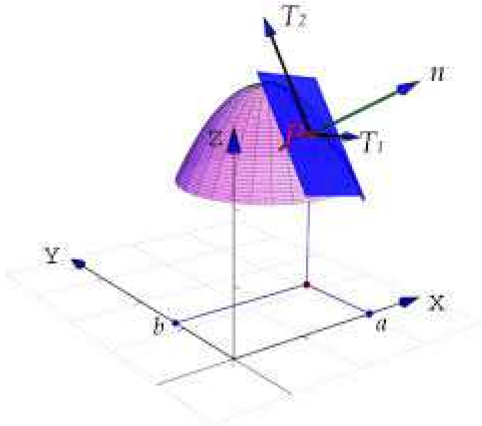


Por lo tanto, el plano tangente deberá contener a estas rectas tangentes, y como además debe pasar por P, la ecuación del plano tangente, si existe, en ese punto está dada por:

$$z = \frac{\partial f\,(a,b)}{\partial x}(x-a) + \frac{\partial f\,(a,b)}{\partial y}(y-b) + f\,(a,b)$$

Nos queda definir el concepto de diferenciabilidad, y lo haremos de manera tal que el plano dado por la ecuación anterior sea una "buena aproximación" a la gráfica de f cerca del punto P, cuando f sea diferenciable.

[1] la ecuación del plano que pasa por un punto (x_0, y_0, z_0) es: $\Pi : A(x - x_0) + B(y - y_0) + C(z - z_0)$



Definición 2.6. *Sea $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ y sea $(x_0, y_0) \in D$. Se dice que f es* ***diferenciable*** *en $(x_0, y_0) \in D$ si:*

$$\frac{f(x,y) - \left\{\frac{\partial f(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y-y_0) + f(x_0,y_0)\right\}}{\sqrt{(x-x_0)^2+(y-y_0)^2}} \to 0$$

cuando (x, y) tiende (x_0, y_0)

Entonces, si se cumple ese límite, el plano tangente

$$\frac{\partial f\,(x_0,y_0)}{\partial x}(x-x_0) + \frac{\partial f\,(x_0,y_0)}{\partial y}(y-y_0) + f\,(x_0,y_0)$$

es una *buena aproximación* a la función $f\,(x, y)$ cuando (x, y) se acerca al punto (x_0, y_0).

Ejemplo 2.7. *Veamos el mecanismo con un ejemplo sencillo.*

Analicemos la diferenciabilidad de $f\,(x, y) = x^2 + y^2$ en el punto $(0, 0)$.

Como es polinómica sabemos que es continua en todos los puntos, buscamos sus derivadas parciales por regla,

$\frac{\partial f(x,y)}{\partial x} = \frac{\partial (x^2+y^2)}{\partial x} = 2x$ *y* $\frac{\partial f(x,y)}{\partial x} = \frac{\partial (x^2+y^2)}{\partial y} = 2y$.

Luego, en el punto $(0, 0)$ ambas dan 0.

Ahora armamos el límite:

$$\frac{f(x,x) - \left\{\frac{\partial f(0,0)}{\partial x}(x-0) + \frac{\partial f(0,0)}{\partial y}(y-0) + f(x_0,y_0)\right\}}{\sqrt{(x-0)^2+(y-0)^2}}$$

$$\frac{(x^2+y^2) - 0x - 0y - 0}{\sqrt{x^2+y^2}} \to 0$$

De esta forma, como el límite tiende a 0 *podemos afirmar que la función es diferenciable en el punto* $(0,0)$.

De manera similar se prueba que la función es diferenciable en todo punto de $\mathbb{R}^2$.

Tenemos algunos resultados teóricos, que aceptaremos sin demostración, que nos ayudan a analizar la diferenciabilidad de las funciones sin recurrir al límite de la definición que a veces puede resultar algo tedioso.

Proposición 2.8. *Si* $f(x,y)$ *es diferenciable en* (x_0,y_0), *entonces* f *es continua en* (x_0,y_0).

En general resulta más útilizar la contrarrecíproca de la afirmación anterior:

Si $f(x,y)$ *NO es continua* en (x_0,y_0), entonces f *NO es diferenciable* en (x_0,y_0).

Teorema 2.9. ***Condición suficiente para la diferenciabilidad***

Sea $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ *y sea* $(x_0, y_0) \in D$*. Si existen las derivadas parciales* $\frac{\partial f(x,y)}{\partial x}$ *y* $\frac{\partial f(x,y)}{\partial y}$*, y además éstas son continuas en un entorno del punto* (x_0, y_0)*, entonces* f *es diferenciable en* (x_0, y_0)*.*

Es importante resaltar que **NO alcanza** con que existan las derivadas parciales de f en el punto dado para que sea diferenciable.
El concepto de diferenciabilidad es "más fuerte" que el de derivabilidad.

En un ejemplo anterior vimos que $f(x, y) = x^2 + y^2$ es una función diferenciable usando la definición, ahora veremos que se comprueba muy fácilmente usando el teorema:

Buscamos las derivadas parciales:

$\frac{\partial f(x,y)}{\partial x} = 2x$ y $\frac{\partial f(x,y)}{\partial x} = 2y$.

Como ambas son continuas (en todo punto de $\mathbb{R}^2$) por el teorema f será diferenciable.

Hay varias propiedades que podemos usar para probar la diferenciabilidad de funciones

Propiedades 2.10. *Dadas* $f : D \subset \mathbb{R}^2 \to R$ *y* $g : D \subset \mathbb{R}^2 \to R$ *funciones diferenciables en un entorno de* $(x_0, y_0) \in D$ *entonces en ese entorno vale que:*

- $f + g$ *es diferenciable*
- $c.f$ *es diferenciable*
- $f.g$ *es diferenciable*
- f/g *es diferenciable (suponemos que* g *nunca es* 0 *en* D*)*

Ejemplo 2.11. *Analizar la diferenciabilidad de:* $f(x,y) = \left\{ \begin{array}{ccc} \frac{xy^2}{x^2+y^2} & si & (x,y) \neq (0,0) \\ 0 & si & (x,y) = (0,0) \end{array} \right\}$

Como tanto el numerador como el denominador son funciones polinómicas (aplicando las propiedades anteriores podemos afirmar que todo polinomio es diferenciable) f es diferenciable en todos los puntos de $\mathbb{R}^2 - (0,0)$.

Falta ver que pasa en el $(0,0)$ y vamos a analizarlo por usando la definición de diferenciabilidad:

Recordemos que f es continua en en $(0,0)$; ahora busquemos las derivadas parciales, también deberemos aplicar la definición,

- $\frac{\partial f(0,0)}{\partial x} = \lim_{h \to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h \to 0} \frac{\frac{h0^2}{h^2+0^2}-0}{h} = 0$

- $\frac{\partial f(0,0)}{\partial y} = \lim_{k \to 0} \frac{f(0,0+k)-f(0,0)}{k} = \lim_{k \to 0} \frac{\frac{0.k^2}{0^2+k^2}-0}{k} = 0$

$$\frac{f(x,y) - \left\{ \frac{\partial f(0,0)}{\partial x}(x-0) + \frac{\partial f(0,0)}{\partial y}(y-0) + f(0,0) \right\}}{\sqrt{(x-0)^2+(y-0)^2}} = \frac{\frac{xy^2}{x^2+y^2} - 0x - 0y - 0}{\sqrt{x^2+y^2}} = \frac{xy^2}{(x^2+y^2)\sqrt{x^2+y^2}} \nrightarrow 0$$

Se puede demostrar que éste límite no tiende a 0 (acercándonos por la trayectoria $y = mx$ por ejemplo), por lo tanto la función NO es diferenciable en $(0,0)$.

Plano Tangente

Ahora que ya definimos el concepto de diferenciabilidad, podemos dar la definición precisa de *plano tangente a la gráfica de una función diferenciable*

Definición 2.12. *Sea $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable en $(x_0, y_0) \in D$. Una ecuación del plano tangente a la gráfica de f en el punto $P_0 = (x_0, y_0, f(x_0, y_0))$ es :*

$$\Pi_T : z = \frac{\partial f\,(x_0, y_0)}{\partial x}\,(x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y}\,(y - y_0) + f\,(x_0, y_0)$$

Ejemplo 2.13. *Dada la función $f(x, y) = x^2 y$, hallar una ecuación para el plano tangente a la gráfica de f por el punto $(-1, 2, f(-1, 2))$*

Antes ya calculamos las derivadas parciales y podemos ver que son funciones continuas, por lo tanto f es diferenciable.

Para hallar la ecuación del plano tangente debemos evaluar la función y sus derivadas parciales primeras en $(-1, 2)$. Entonces:

$$\begin{array}{rcl} f\,(-1,2) & = & (-1)^2 2 = 2 \\ \dfrac{\partial f\,(x,y)}{\partial x} & = & 2xy \to \dfrac{\partial f\,(-1,2)}{\partial x} = 2 \cdot (-1).2 = -4 \\ \dfrac{\partial f\,(x,y)}{\partial y} & = & x^2 \to \dfrac{\partial f\,(-1,2)}{\partial x} = (-1)^2 = 1 \end{array}$$

Por lo tanto, el plano tangente a la gráfica de f por el punto $(-1, 2, f(-1, 2))$ es:

$$\Pi_T : z = 2 + (-4)(x - (-1)) + 1(y - 2) = 2 - 4(x + 1) + (y - 2) = -4 - 4x + y$$

Linealización

Se denomina linealización de f en el punto $P_0 = (x_0, y_0)$, o también polinomio de Taylor de primer orden alrededor de P_0, a la siguiente función de dos variables:

$$L(x,y) = f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y - y_0)$$

cuya gráfica es el plano tangente Π_T a f en el punto P_0.

De manera que si f es diferenciable en un punto P_0 de su dominio, entonces su gráfica admite un plano tangente en el punto ese punto, además este plano es localmente una "buena aproximación" a la gráfica de f cerca de P_0, y por tanto podemos usar la función linealización para calcular la ***aproximación lineal*** de f en P_0:

$$f(x,y) \approx L(x,y) = f(x_0,y_0) + \frac{\partial f(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial f(x_0,y_0)}{\partial y}(y - y_0)$$

Ejemplo 2.14. *Considerar la función* $f(x,y) = e^{x^2} + e^{y^2}$,
Hallar la linealización de f *en* $(-1,1)$ *y la aproximación lineal para* f *en* $(1,0)$.

Para encontrar la linealización debemos encontrar las derivadas parciales en el punto pedido y además asegurarnos que la función sea diferenciable en ese punto del dominio.
Claramente nuestra f *es una función continuas con derivadas continuas.*

$*$ $\frac{\partial f(x,y)}{\partial x} = \frac{\partial e^{x^2}+e^{y^2}}{\partial x} = e^{x^2}2x + 0$ *y en el punto* $\frac{\partial f(-1,1)}{\partial x} = e^{(-1)^2}2(-1) = -2e$

$*$ $\frac{\partial f(x,y)}{\partial y} = \frac{\partial e^{x^2}+e^{y^2}}{\partial y} = 0 + e^{x^2}2y$ *y en el punto* $\frac{\partial f(-1,1)}{\partial y} = e^{(1)^2}2(1) = 2e$

Además, $f(-1,1) = e^{(-1)^2} + e^{1^2} = e^1 + e^1 = e + e = 2e$

Luego, La linealización de f *alrededor de* $(-1,1)$ *está dada por la función:*

$$L(x,y) = 2e - 2e(x+1) + 2e(y-1)$$

Usemos la función de linealización para aproximar f en el punto $(-1.1, 0.9)$:

$f(-1.1, 0.9) \approx L(-1.1, 0.9) = 2e - 2e(-1.1+1) + 2e(0.9-1) = 2e - 2e(-0.1) + 2e(-0.1) = 2e + 2e(0.1) - 2e(0.1) = 2e$

(Si calcucamos el valor ”'exacto”' de $f(-1.1, 0.9)$ veremos que da $e^{(-1.1)^2} + e^{(0.9)^2} = 5.6013926392...$ que es cercano a 2e)

Calculemos ahora la aproximación lineal de f en $(0, 1)$ está dada por la expresión:

$L(x, y) = f(1, 0) + \frac{\partial f(1,0)}{\partial x}(x-1) + \frac{\partial f(1,0)}{\partial y}(y-0) = 2e - 2e(x+1) + 2e(y-1)$

entonces $f(0, 1) \approx L(0, 1) = 2e - 2e(0+1) + 2e(1-1) = 2e - 2e.1 + 2e.0 = 0$
PERO $f(0, 1) = e^{0^2} + e^{1^2} = 1 + e$ *no se aproxima a 0!!!*

Esto nos pasa por buscar la aproximación en un punto no cercano al $(-1, 1)$

2.3 Vector Gradiente y Derivada Direccional

El **vector gradiente** de una función de dos variables $f(x, y)$ es el vector cuyas componentes son las derivadas parciales de la función f, se lo indica ∇f.

Esto es: $\nabla f(x, y) = \left(\frac{\partial f(x,y)}{\partial x} \frac{\partial f(x,y)}{\partial y}\right) = (f_x(x, y), f_y(x, y))$

Ejemplos 2.15. *Veamos algunos ejemplos:*

- *El vector gradiente de la función $f(x, y) = x^2y$ es:*
  $\nabla f(x, y) = \left(\frac{\partial f(x,y)}{\partial x} \frac{\partial f(x,y)}{\partial y}\right) = \left(\frac{\partial x^2y}{\partial x}, \frac{\partial x^2y}{\partial y}\right) = (2xy, x^2)$
- *El gradiente de la función $g(x, y) = e^{x^2} + e^{y^2}$ es:* $\nabla g(x, y) = \left(\frac{\partial e^{x^2}+e^{y^2}}{\partial x}, \frac{\partial e^{x^2}+e^{y^2}}{\partial y}\right) = (2xe^{x^2}, 2ye^{y^2})$

Vamos a usar el vector gradiente para estudiar la variación de la función $f(x,y)$ en una dirección cualquiera.

Así como en una variable la derivada de una dada función representaba la dirección de crecimiento o decrecimiento, en varias variables es la derivada direccional en la dirección de un vector dado la que representa la tasa de cambio de la función en la dirección de ese vector.

Definición 2.16. *Sea $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ la derivada direccional de f en $\boldsymbol{x}$ en la dirección del vector unitario $\vec{v} = (v_1, v_2)$ está dada por:*

$$D_{\vec{v}} f(x,y) = \lim_{t \to 0} \frac{f\left((x,y) + t(v_1, v_2)\right) - f(x,y)}{t}$$

si es que el límite existe.

Teorema 2.17. *Si $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ es diferenciable, entonces existen todas las derivadas direccionales. La derivada direccional de f en $\boldsymbol{x}$ en la dirección de $\vec{v}$ está dada por:*

$$Df\left(\boldsymbol{x}\right) = \nabla f\left(\boldsymbol{x}\right) \cdot \vec{v}$$

donde $\nabla f\left(\boldsymbol{x}\right) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$

Ejemplo 2.18. *Calculemos la derivada direccional de $f(x,y) = x^2.y$ en la dirección del vector $\vec{v} = (1,1)$ en el punto $(-1,3)$) (en este caso lo haremos por ambos métodos ya que al ser f diferenciable podemos aplicar el teorema).*

El vector unitario en la dirección de $\vec{v}$ es $\vec{u} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$

- *Por definición*
  $D_{\vec{v}} f(x,y) = \lim_{t\to 0} \frac{f((x,y)+t(v_1,v_2))-f(x,y)}{t} = \lim_{t\to 0} \frac{f\left((-1,3)+t(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})\right)-f(-1,3)}{t}$

  $= \lim_{t\to 0} \frac{f(-1+t\frac{1}{\sqrt{2}},3+t\frac{1}{\sqrt{2}}))-3}{t} = \lim_{t\to 0} \frac{(-1+t\frac{1}{\sqrt{2}})^2(3+t\frac{1}{\sqrt{2}})-3}{t} = \lim_{t\to 0} \frac{(1-\frac{2t}{\sqrt{2}}+\frac{t^2}{2})(3+t\frac{1}{\sqrt{2}})-3}{t}$

  $= \lim_{t\to 0} \frac{3-\frac{6t}{\sqrt{2}}+\frac{3t^2}{2}+t\frac{1}{\sqrt{2}}-2t^2\frac{1}{\sqrt{2}\sqrt{2}}+t^3\frac{1}{2\sqrt{2}}-3}{t} = \lim_{t\to 0} \frac{-\frac{5t}{\sqrt{2}}-\frac{t^2}{2}+t^3\frac{1}{2\sqrt{2}}}{t} = -\frac{5}{\sqrt{2}}$

- *Usando el teorema* $D_{\vec{u}}f(a,b) = \nabla f(a,b) \cdot \vec{u} = \nabla f(-1,3) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \big(2(-1)3, (-1)^2\big) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = (-6,1) \cdot (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \frac{-6}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \frac{-5}{\sqrt{2}}$

Dirección de máximo crecimiento

Recordando propiedades del producto escalar entre vectores, sabemos que $v \cdot w = |v|\,|w|\,cos(\theta)$ (siendo θ el ángulo entre los dos vectores), la derivada direccional puede escribirse, usando que $\vec{u} = 1$ como

$$D_{\vec{u}}f(a,b) = |\nabla f(a,b)|\,|\vec{u}|\,cos(\theta) = |\nabla f(a,b)|\,cos(\theta)$$

De la expresión anterior deducimos que la derivada direccional tendrá un valor máximo cuando $cos(\theta) = 1$, lo que ocurre cuando $\theta = 0$, es decir cuando $\vec{u}$ y $\nabla f(a,b)$ tienen la misma dirección.

En tal caso, tendremos:

$$D_{\vec{u}}f(a,b) = |\nabla f(a,b)|$$

Teorema 2.19. *Si $f(x,y)$ es diferenciable en (ab) la **dirección de máximo crecimiento** de f en (a,b) está dada por la dirección del gradiente de f en (a,b).*

Además, la máxima razón de cambio *es* $|\nabla f(a,b)|$

Ejemplo 2.20. *Encontrar la dirección de máximo crecimiento y las razones de cambio máxima y mínima de la función $g(x,y) = e^{x^2} + e^{y^2}$ en el punto $(1,2)$*

Como nuestra función es diferenciable la dirección de máximo crecimiento está dada por el vector gradiente ∇g en el punto. Lo calculamos:

$\nabla g(x,y) = \left(2xe^{x^2}, 2ye^{y^2}\right)$, *por lo tanto* $\nabla g(1,2) = \left(2.1e^{1^2}, 2.2e^{2^2}\right) = \left(2e^2, 4e^4\right)$ *es la dirección de máximo crecimiento.*

La razón de cambio máxima será: $|\nabla g(1,2)| = \left|\left(2e^2, 4e^4\right)\right| = \sqrt{4e^4 + 16e^{16}}$ *y obviamente la razón de cambio mínima es:* $-\sqrt{4e^4 + 16e^{16}}$

3 Extremos de funciones de dos variables

Vamos a definir la noción de extremos en funciones de dos variables (de manera similar se puede definir para más variables).

Definición 3.1. *Sea $f : D \subset R^2 \to R$. Se dice que f tiene un* ***máximo local o relativo en*** $(x_0, y_0) \in D$*, si* $f(x_0, y_0) \geq f(x, y)$ *para todo punto* (x, y) *en algún disco centrado en* (x_0, y_0).
Si la desigualdad se verifica para todo punto del dominio de f , el máximo es ***absoluto o global.***

Se dice que f tiene un ***mínimo local o relativo en*** $(x_0, y_0) \in D$ *si* $f(x_0, y_0) \leq f(x, y)$ *para todo punto* (x, y) *en algún disco centrado en* (x_0, y_0).
Si la desigualdad se verifica para todo punto del dominio de f , el mínimo es ***absoluto o global.***

Los máximos y mínimos locales constituyen los ***extremos locales o relativos*** *de f .*

Esto dice que si f tiene un máximo relativo en un punto, entonces el valor de la función en dicho punto es el mayor valor que toma f para cualquier punto en los alrededores; mientras que si f tiene un mínimo relativo,la función en dicho punto es el menor valor que toma f.

Al igual que en Matemática 2 vamos a utilizar el concepto de punto crítico:

Definición 3.2. *Sea $f : D \subset R^2 \to R$. Un punto (x_0, y_0) perteneciente al dominio de f es un* ***punto crítico*** *de f si se cumplen algunas de las siguientes condiciones:*

- *Las dos derivadas parciales primeras de f se anulan en (x_0, y_0)*
- *Al menos una de las derivadas parciales primeras de f* ***no*** *existe en (x_0, y_0)*

En el primer caso (o sea, cuando el gradiente de f en el punto es el vector nulo), se dice que (x_0, y_0) es un ***punto estacionario*** *de f .*

El siguiente teorema, que daremos sin demostración, afirma que todo extremo local es un punto crítico pero no todo punto crítico es un extremo.

Teorema 3.3. *Sea $f : D \subset R^2 \to R$, D un conjunto abierto.*

Si f tiene un máximo o mínimo local en $(x_0, y_0) \in D$ y si existen las derivadas parciales primeras de f en dicho punto, entonces $f_x(x_0, y_0) = 0$ y $f_y(x_0, y_0) = 0$.

(Equivalentemente, $\nabla f(x_0, y_0) = 0$).

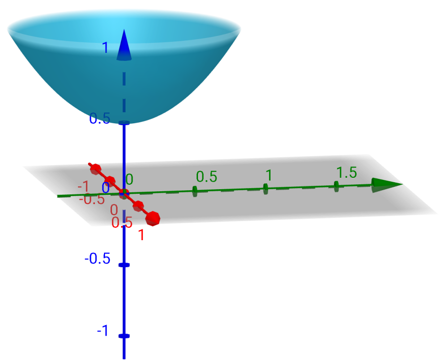
Para hallar extremos de una función, se buscan todos los puntos críticos y luego se analiza si los "candidatos a extremos" realmente lo son.

Ejemplo 3.4. *Calculemos los extremos de $f(x, y) = x^2 + y^2 + 5$*

f es una función polinómica y por lo tanto será diferenciable, entonces sabemos que sus derivadas parciales existen.

$$f_x(x, y) = 2x \qquad f_y(x, y) = 2y$$

Ambas se anularán cuando $x = 0$, $y = 0$, es decir, unicamente en $(x, y) = (0, 0)$
Como $f(0, 0) = 5 \leq x^2 + y^2 + 5 = f(x, y)$ para todo (x, y), hay un m ınimo local en $(0, 0)$.



Teorema 3.5. Teorema de Clairaut

Sea $f : D \to \mathbb{R}$ una función de dos variables definida en un conjunto abierto $D \subseteq \mathbb{R}^2$, si existen las segundas derivadas cruzadas y son continuas en D, entonces estas son iguales, es decir:

$$\frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 f}{\partial y\, \partial x}$$

Teorema 3.6. Criterio de las derivadas segundas

Sea $f : D \to \mathbb{R}$ una función de dos variables definida en un conjunto abierto $D \subseteq \mathbb{R}^2$con segundas derivadas continuas en un entorno de un punto crítico (x_0, y_0), llamemos

$$H = \frac{\partial^2 f}{\partial x\, \partial x}(x_0, y_0)\frac{\partial^2 f}{\partial y\, \partial y}(x_0, y_0) - \left(\frac{\partial^2 f}{\partial x\, \partial y}(x_0, y_0)\right)^2$$

entonces,

- *Si $H > 0$ y $\frac{\partial^2 f}{\partial x; \partial x}(x_0, y_0) > 0$ entonces $f(x, y)$ tiene un mínimo local en (x_0, y_0)*
- *Si $H > 0$ y $\frac{\partial^2 f}{\partial x\, \partial x}(x_0, y_0) < 0$ entonces $f(x, y)$ tiene un máximo local en (x_0, y_0)*
- *Si $H < 0$ entonces $f(x, y)$ no tiene ni un máximo ni un mínimo en (x_0, y_0). Diremos que f tiene un **punto de ensilladura** (o **punto silla**) en (x_0, y_0).*
- *Si $H = 0$ el criterio no decide.*

(Obs: se puede usar $\frac{\partial^2 f}{\partial y\, \partial y}$ en lugar de $\frac{\partial^2 f}{\partial x\, \partial x}$)

3.1 Descenso del Gradiente

Si bien es posible determinar los extremos de una función mediante métodos analíticos, como la búsqueda de puntos estacionarios y su categorización mediante el criterio de las derivadas segundas, existen algunas funciones que no resultan muy fáciles de estudiar. La optimización de funciones de una o múltiples variables, es decir, encontrar sus extremos, es una de las aplicaciones fundamentales del análisis matemático a lo largo de todas las ciencias.

El advenimiento de la informática en ramas de la ciencia como, Matemática y Física, hizo posible la construcción de soluciones algorítmicas a los ya bien conocidos métodos analíticos y algebráicos. El análisis numérico surgió para proveer de una nueva forma automatizada de resolución de problemas matemáticos, entre ellos la aproximación de extremos de funciones complejas de estudiar. Cabe destacar que los resultados y algoritmos que brinda el análisis numérico logran simples aproximaciones de los valores verdaderos, pero estas, en general, pueden acercarse tanto como uno quiera los valores analíticos reales.

En esta sección introduciremos un método numérico de optimización de funciones conocido como **descenso del gradiente** que nos permitira minimizar una función $f(\mathbf{x})$ aproximando de forma iterativa lo mejor posible la solucion analitica verdadera. Resulta importante aclarar que éste es solo uno de los metodos existentes para solucionar este problema de optimizacion. Asi como su nombre lo indica, el gradiente de la función en cuestión estará involucrado en el proceso iterativo de minimización.

Los metodos de descenso, en general, utilizan la misma idea subyacente para conseguir una aproximación de la solución óptima: Partiendo de un punto inicial $x^{(0)}$, producir una secuencia de puntos $x^{(1)}, x^{(2)}, \dots, x^{(k)}$ tales que se minimice cada vez mas la función, es decir,

$$f(x^{(0)}) \geq f(x^{(1)}) \geq f(x^{(2)}) \geq \cdots \geq f(x^{(k)})$$

Suponiendo siempre que $x^{(0)}$ no es el punto donde la funcion consigue su valor minimo.

Para construir dicha secuencia de puntos partiendo de uno inicial debemos decidir, en cierta forma, para dónde debemos trasladarnos. Podriamos movernos sobre una recta (en $\mathbb{R}^2$ cuando la funcion es de una variable) o sobre un plano (en $\mathbb{R}^3$ cuando la funcion es de dos variables).

Comencemos interpretando nuestro problema de forma simbólica: buscamos construir una secuencia de puntos $x^{(t+1)}$ que, tomando su anterior $x^{(t)}$, lo modifique de forma tal que en cada paso nos acerquemos un poco más a la solución mínima. Dicha modificación la realizaremos utilizando un vector, que funcionará como una brújula y se ajustará a cada paso para indicar nuestro camino hacia el "norte" (la solución óptima). Llamando Δx a dicho vector indicador, veamos que podemos expresar a cada nuevo punto como una actualización del anterior de la siguiente forma:

$$x^{(t+1)} = x^{(t)} + \eta \Delta x^{(t)} \tag{2}$$

Como sambemos, si sumamos un vector u a uno v, obtenemos otro vector s que tendrá la punta de v "trasladada" $|u|$ unidades en la direccion de u. Teniendo esto en cuenta, podemos ver que la ecuación (2) indica una traslación del punto $x^{(t)}$ en la dirección de nuestro vector "brújula" $\Delta x^{(t)}$. El factor η (eta) indica un escalamiento del vector de direccionamiento y lo llamaremos *tamaño de paso* (también conocido como *tasa de aprendizaje* en el contexto del Deep Learning). Veremos que este factor será necesario para conseguir una aproximación correcta a medida que nos acercamos a la solución verdadera. No siempre querremos que nuestro punto se actualice trasladandose la mismas unidades en cada paso, menos aun cuando nos estemos acercando al punto mínimo, donde buscaremos realizar modificaciones pequeñas, para asegurarnos presición y no "pasarnos de largo".

Luego, como el análisis nos ha enseñado, el vector gradiente nos indica la dirección de máximo o mínimo crecimiento de la función, y éste será nada más y nada menos que el vector "brújula" $\Delta x^{(t)}$. De esta forma, como $-\nabla f(\mathbf{x})$ indica la dirección de mínimo crecimiento de la función f en el punto $\mathbf{x}$, la ecuación (2) pasa a ser:

$$x^{(t+1)} = x^{(t)} - \eta \nabla f(x^{(t)}) \tag{3}$$

Tomando (3) , construiremos tantos puntos de forma tal que las actualizaciones cercanas a la solucion verdadera sean cada vez menores. Pero... ¿cuantas debemos generar?

Notemos que cuando la función tome valores muy similares en dos aproximaciones consecutivas, como trataremos con funciones continuas, seria sensato pensar que la actualizacion fue muy pequeña y por ende nos encontramos cerca de la solución. Esta **condición de corte** tendrá la forma:

$$|f(x^{(k+1)}) - f(x^{(k)})| < \varepsilon$$

donde ε es un valor muy pequeño, en el orden de 10^{-5} o menor (cuanto menor mejor será la aproximación, pero mayor cantidad de iteraciones serán requeridas).

3.1.1 Construcción del Algoritmo

Es hora de construir un paso a paso del **descenso del gradiente**. Buscamos que, a partir de un punto inicial dado x_0, se construya una secuencia de puntos nuevos $x_1, \ldots, x_n$ de forma iterativa que minimice la función f a cada paso. Pasemos entonces a construir el algoritmo:

1. Seleccionar un punto inicial x_0
2. Mientras no se haya llegado a una buena aproximación del mínimo la función:
   - Calcular el siguiente punto de la secuencia según la función de actualización:

$$x^{(t+1)} = x^{(t)} - \eta \nabla f(x^{(t)})$$

Pensando más en detalle, el punto 2 de nuestro problema implica definir una condición de corte. Esta será la misma que mencionamos previamente. Generaremos una secuencia de puntos hasta que dos consecutivos den un valor en la función muy similar, es decir, hasta que se cumpla:

$$|f(x^{(k+1)}) - f(x^{(k)})| < \varepsilon$$

El parámetro ε lo llamaremos **tolerancia** y será necesario seleccionarlo de forma previa. Así mismo es necesario definir el valor de tamaño de paso η. Por lo tanto, nuestro algorimo, quedría finalemente de la forma:

1. Seleccionar un punto inicial x_0
2. Seleccionar el valor de tolerancia ε
3. Seleccionar el tamaño del paso η
4. Mientras no se cumpla la condición de corte: $|f(x^{(k+1)}) - f(x^{(k)})| < \varepsilon$
   - Calcular el siguiente punto de la secuencia según la función de actualización:

$$x^{(t+1)} = x^{(t)} - \eta \nabla f(x^{(t)})$$

Ejemplo 3.7. *Veamos a continuación un ejemplo del algoritmo paso a paso. Por simplicidad y explicabilidad, tomaremos una función de una variable:*

$$f(x) = \frac{1}{2}x^2 + \frac{3}{2}$$

De Matematica 2, sabemos que esta función tiene un único mínimo global y es fácilmente hallado mediante el cálculo de la derivada:

$$f'(x) = \frac{d}{dx}\left(\frac{1}{2}x^2 + \frac{3}{2}\right) = x$$

Igualando $f'(x)$ a 0 podemos hallar el punto crítico que resulta ser donde la función alcanza su mínimo. Luego, tenemos que en $x = 0$ se encuentra el mínimo de la función.

Comencemos el algoritmo:

- *Seleccionamos el punto inicial $x_0 = 1.1$*
- *Seleccionamos $\varepsilon = 10^{-2} = 0.01$ y $\eta = 0.4$*
- ***Iteración 1:***
  - *Como es la primera iteración, calculamos el siguiente punto x_1*

    $$x_1 = x_0 - \eta\,\nabla f(x_0) = 1.1 - 0.4f'(1.1) = 1.1 - (0.4\,.\,1.1) = 0.66$$

  - *Verificamos la condición de corte:*

    $$|f(x_1) - f(x_0)| = |f(0.66) - f(1.1)| = |1.718 - 2.105| = 0.387$$

    Como $0.387 > 0.01$ debemos continuar el algoritmo.
- ***Iteración 2:***
  - *Calculamos el siguiente punto x_2*

    $$x_2 = x_1 - \eta\,\nabla f(x_1) = 0.66 - 0.4f'(0.66) = 0.66 - (0.4\,.\,0.66) = 0.396$$

  - *Verificamos la condición de corte:*

    $$|f(x_2) - f(x_1)| = |f(0.396) - f(0.66)| = |1.578 - 1.718| = 0.14$$

    Como $0.14 > 0.01$ debemos continuar el algoritmo.

- ***Iteración 3:***
  - *Calculamos el siguiente punto* x_3

$$x_3 = x_2 - \eta \nabla f(x_2) = 0.396 - 0.4f'(0.396) = 0.396 - (0.4\,.\,0.396) = 0.237$$

  - *Verificamos la condición de corte:*

$$|f(x_3) - f(x_2)| = |f(0.237) - f(0.396)| = |1.528 - 1.578| = 0.05$$

    Como $0.14 > 0.01$ *debemos continuar el algoritmo.*

- ***Iteración 4:***
  - *Calculamos el siguiente punto* x_4

$$x_4 = x_3 - \eta \nabla f(x_3) = 0.237 - 0.4f'(0.237) = 0.237 - (0.4\,.\,0.237) = 0.142$$

  - *Verificamos la condición de corte:*

$$|f(x_4) - f(x_3)| = |f(0.142) - f(0.237)| = |1.510 - 1.528| = 0.018$$

    Como $0.018 > 0.01$ *debemos continuar el algoritmo.*

- ***Iteración 5:***
  - *Calculamos el siguiente punto* x_4

$$x_5 = x_4 - \eta \nabla f(x_4) = 0.142 - 0.4f'(0.142) = 0.142 - (0.4\,.\,0.142) = 0.085$$

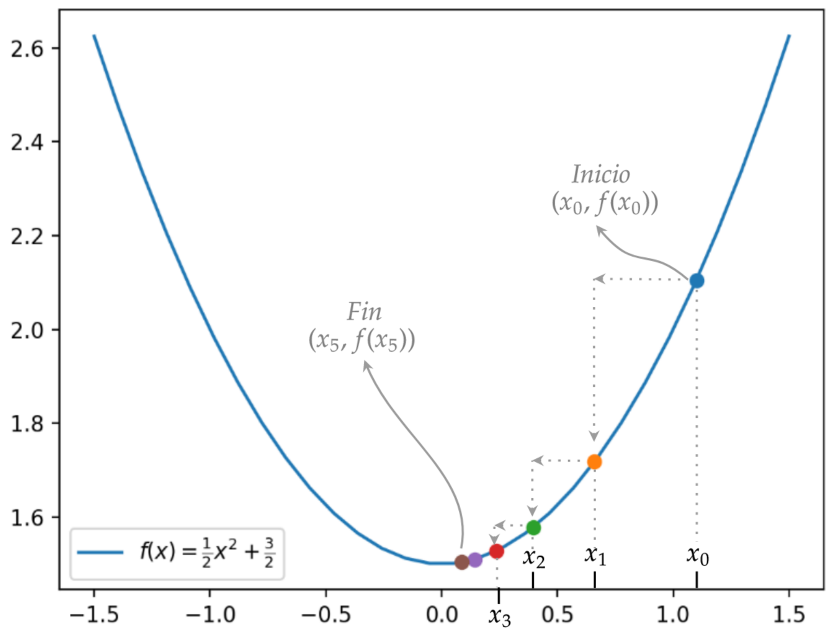
  - *Verificamos la condición de corte:*

$$|f(x_5) - f(x_4)| = |f(0.085) - f(0.142)| = |1.503 - 1.510| = 0.007$$

    Finalmente, como $0.007 < 0.01$ *detenemos la ejecución del algoritmo.*

  Por lo tanto, obtenemos que el mínimo aproximado de la función es $x* = 0.085$ *que resulta bastance cercano al verdadero* $x = 0$. *Cabe destacar que si eligieramos una tolerancia mucho menor, por ejemplo, en el orden de* 10^{-5}, *obtendríamos una aproximacion mucho mas precisa.*

Por último veamos de forma gráfica lo que estuvimos realizando (pueden acceder al código que lo generó con este enlace):



iteration	prev	new	f(new)	error
1	1.10000	0.66000	1.717800	0.387200
2	0.66000	0.396000	1.578408	0.139392
3	0.396000	0.237600	1.528227	0.050181
4	0.237600	0.142560	1.510162	0.018065
5	0.142560	0.085536	1.503658	0.006503