Métodos III - Derivadas

February 12, 2019

5 Gradiente, matriz jacobiana, funciones diferenciables

Jose A. Hernando
Departamento de Física de Partículas. Universidade de Santiago de Compostela
Enero 2019

In [1]:
```
import time
print(' Last version ', time.asctime() )
```

```
 Last version  Tue Feb 12 20:32:35 2019
```

5.1 Objectivos

Presentar los conceptos de

- gradiente y matriz jacobiana
- definición de función diferenciable

Discutir sobre:

- La condición suficiente de diferenciabilidad
- las bondades de las funciones diferenciables

In [2]:
```
# general imports
# general imports
%matplotlib inline
%reload_ext autoreload
%autoreload 2

# numpy and matplotlib
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
matplotlib.style.use('ggplot')
import graph_utils as gf

figsize = 6, 3.8
cmap    = 'hot'
```

5.2 Gradiente

Para una función escalar $f(x,y)$ diferenciable, hemos dado el desarrollo de Taylor en un punto (x_0, y_0) si nos desplazamos en el espacio origen un vector (v_x, v_y).

$$f((x_0,y_0)+(v_x,v_y)) = f(x_0,y_0) + f'_x(x_0,y_0)v_x + f'_y(x_0,y_0)v_y$$

o lo que es lo mismo:

$$f((x_0,y_0)+(v_x,v_y)) = f(x_0,y_0) + \Big(f'_x(x_0,y_0),\, f'_y(x_0,y_0)\Big)\ (v_x,v_y)$$

Donde la expresión:

$$\Big(f'_x(x_0,y_0),\, f'_y(x_0,y_0)\Big)\ (v_x,v_y)$$

Es el producto vectorial entre (v_x, v_y), el vector desplazamiento; y otro vector, el **gradiente**, $\nabla f(x_0, y_0)$, cuyas componentes son las derivadas parciales:

$$\nabla f(x_0,y_0) = \Big(f'_x(x_0,y_0),\, f'_y(x_0,y_0)\Big) = \left(\frac{\partial f(x_0,y_0)}{\partial x}, \frac{f(x_0,y_0)}{\partial y}\right)$$

Supongo que ya has visto el gradiente en Física. En el electromagnetismo definimos el campo eléctrico como (menos) el gradiente del potencial.

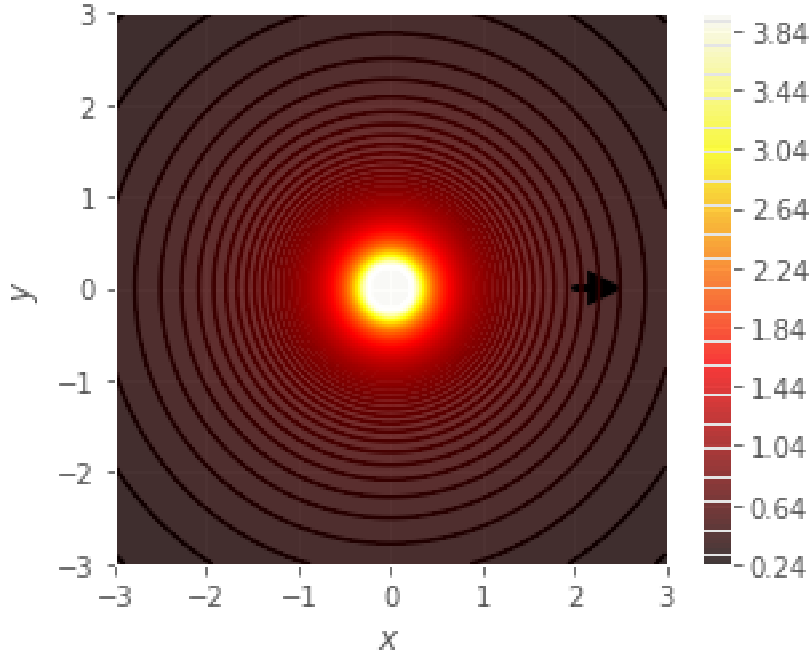
$$\mathbf{E}(\mathbf{x}) = -\nabla V(\mathbf{x})$$

Puedes revisar que efectivamente es así, a partir de la definición de potencial y campo eléctrico que vimos en la sección sobre funciones escalares y vectoriales.

Ejemplo: La siguiente figura nos muestra los conjuntos de nivel del potencial eléctrico de una carga. También nos muestra el campo eléctrico, (el gradiente cambiado de signo), en un determinado punto (x_0, y_0)

Explora y cambia el punto y observa como cambia el gradiente.

```
In [3]: x0, y0 =  2, 0.
        V  = lambda x, y : 1/(x*x + y*y)**(1/2)
        Ex = lambda x, y : x/(x*x + y*y)**(3/2)
        Ey = lambda x, y : y/(x*x + y*y)**(3/2)
        gf.contour(V, contours = 100, zlim=(0., 4));
        gf.arrow(x0, y0, Ex(x0, y0), Ey(x0, y0));
```



Por supuesto, lo que hemos visto para funciones escalares de dos dimensiones para para cualquier dimensión.

Sea una función $f(\mathbf{x})$ escalar definida en R^n, diremos que es "suave", si tiene desarrollo de Taylor de primer orden, esto es, la función en un punto $\mathbf{x}$ próximo a $\mathbf{x}_0$, relacionados por un vector "pequeño", $\mathbf{v}$ por $\mathbf{x} = \mathbf{x}_0 + \mathbf{v}$, se puede aproximar por un hyper-plano.

$$f(\mathbf{x}_0 + \mathbf{v}) = f(\mathbf{x_0}) + \nabla f(\mathbf{x_0})\,\mathbf{v}$$

donde las derivadas parciales son las pendientes en cada dirección, $f'_i(\mathbf{x}_0)$.

Ejercicio: Calcula el gradiente de las siguientes funciones:

1. $f(x,y) = e^{x+y}$
2. $f(x,y) = \cos x \sin y$

Solución:

1. $\nabla f(x,y) = (e^{x+y},\ e^{x+y})$
2. $\nabla f(x,y) = (-\sin x \cos y,\ \cos x \cos y)$

El gradiente nos aporta más información sobre la función. Fíjate:

El término:

$$\nabla f(\mathbf{x}_0)\,\mathbf{v}$$

es un producto escalar que podemos reescribir cómo:

$$\nabla f(\mathbf{x}_0)\,\mathbf{v} = ||\nabla f(\mathbf{x}_0)||\,||\mathbf{v}||\,\cos\theta$$

donde θ es el ángulo que forman los dos vectores.

Tenemos ahora dos casos:

1. los dos vectores son paralelos, si van en la misma dirección y sentido, es máximo.

Esto es, si el desplazamiento **v** va en la misma dirección y sentido que el gradiente, el cambio de la función es máximo.

2. los dos vectores son perpendiculares, el término es nulo.

O lo que es lo mismo, la valor función no cambiará ortogonalmente al gradiente.

Regresemos a la relación entre el campo y el potencial eléctrico:

$$\mathbf{E}(\mathbf{x}) = -\nabla V(\mathbf{x})$$

El gradiente entonces indica la dirección y sentido en el que el potencial eléctrico *decae* más rápidamente.

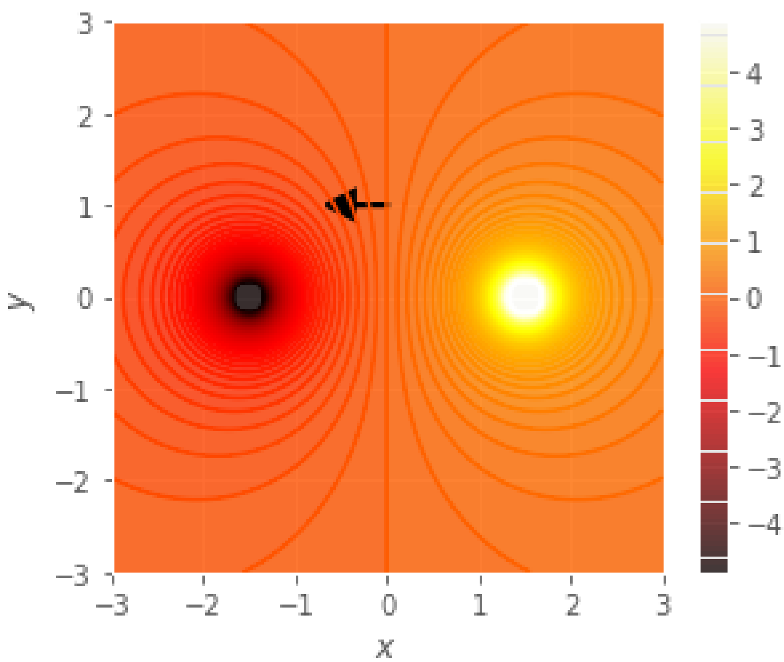
Ahora recordemos los conjuntos de nivel, que son aquellos puntos del espacio inicial cuya valor de la función es c.

El gradiente va a ser normal a los conjuntos de nivel, porque en los conjuntos de nivel el valor de la función no cambia.

Ejemplo: La siguiente figura nos muestra los conjuntos de nivel de un dipolo eléctrico. También nos muestra el campo eléctrico, (el gradiente cambiado de signo), en un determinado punto (x_0, y_0)

Explora y cambia el punto y observa como cambia el gradiente.

```
In [4]: from common_functions import edipole
        x0, y0 = 0, 1.
        V, Ex, Ey = edipole(1., 1.5, 0., -1.5, 0.)
        gf.contour(V, contours= 100, zlim=(-5., 5.))
        gf.arrow(x0, y0, Ex(x0, y0), Ey(x0, y0));
```

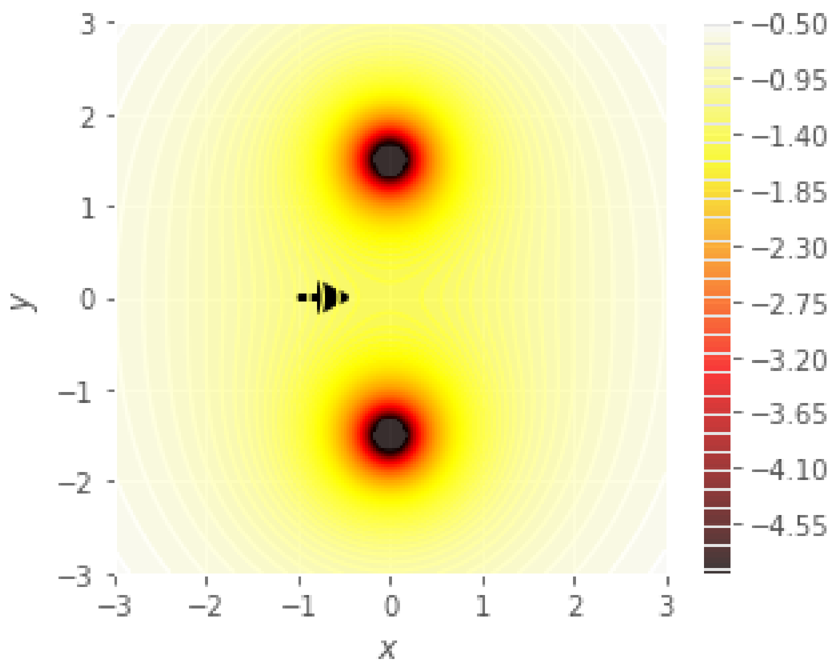


La siguiente celda te permite dibujar el potencial de un conjunto de cargas. El dibujo corresponde a dos cargas negativas situadas en el eje vertical separadas por 2 unidades.

Explora y cambia el punto y observa como cambia el potencial y el gradiente.

Explora y cambia el sistema de cargas para ver como cambia el potencial y el gradiente en un punto.

```
In [5]: from common_functions import esystem
        x0, y0 = -1., 0.
        charges = [(-1., 0, 1.5), (-1., 0, -1.5)]
        V, Ex, Ey = esystem(charges)
        gf.contour(V, contours = 100, zlim=(-5., 5.))
        gf.arrow(x0, y0, Ex(x0, y0), Ey(x0, y0));
```



5.3 Matriz Jacobiana

Una función vectorial es diferenciable, si y solo si, lo son cada una de sus funciones componentes.

Sea una función vectorial, $\mathbf{f}(\mathbf{x})$ de $\mathrm{R}^n \to \mathrm{R}^m$, que es diferenciable en un punto $\mathbf{x}_0$

cada uno de sus funciones componentes debe ser diferenciable, pongamoslas una debajo de la otra:

$$f_1(\mathbf{x}_0 + \mathbf{v}) \approx f_1(\mathbf{x}_0) + \nabla f_1(\mathbf{x}_0)\mathbf{v}$$

$$\ldots$$

$$f_m(\mathbf{x}_0+\mathbf{v}) \approx f_i(\mathbf{x}_0) + \nabla f_m(\mathbf{x}_0)\mathbf{v}$$

O también:

$$f_1(\mathbf{x}_0+\mathbf{v}) \approx f_1(\mathbf{x}_0) + \sum_{i=1,n} \frac{\partial f_1(\mathbf{x}_0)}{\partial x_i} v_i$$
$$\dots$$
$$f_m(\mathbf{x}_0+\mathbf{v}) \approx f_m(\mathbf{x}_0) + \sum_{i=1,n} \frac{\partial f_m(\mathbf{x}_0)}{\partial x_i} v_i$$

Que podemos reescribir de forma matricial, colocando $\mathbf{v}$, $\mathbf{f}(\mathbf{x}_0+\mathbf{v})$ y $\mathbf{f}(\mathbf{x}_0)$ como vectores columnas.

$$\mathbf{f}(\mathbf{x}_0+\mathbf{v}) = f(\mathbf{x}_0) + \mathbf{Df}(\mathbf{x}_0)\,\mathbf{v}$$

Donde $\mathbf{Df}(\mathbf{x}_0)$ es una matriz cuyos elementos son:

$$\mathbf{Df}_{ij}(\mathbf{x}_0) = \frac{\partial f_i(\mathbf{x_0})}{\partial x_j}$$

Cuestión: Si te fijas en las filas de la matriz jacobiana, ¿a qué corresponden?
Ejercicio: Calcula la matriz jacobiana de la siguiente función vectorial:

$$\mathbf{f}(x,y) = \left(e^{x+y}, \cos x \sin y\right)$$

Solución:
$\mathbf{Df}(x,y) = \begin{pmatrix} e^{x+y} & e^{x+y} \\ -\sin x \sin y & \cos x \cos y \end{pmatrix}$
Recapitulemos:

- Si una función escalar, $f(\mathbf{x})$ es diferenciable podemos aproximarla en un punto $\mathbf{x}_0$ y en una dirección $\mathbf{v}$ por :

$$f(\mathbf{x}_0+\mathbf{v}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)\,\mathbf{v}$$

donde $\nabla f \mathbf{x}_0$ es el gradiente:

$$\nabla f(x_0,y_0) = \left(\frac{\partial f(x_0,y_0)}{\partial x}, \frac{f(x_0,y_0)}{\partial y}\right)$$

- Si una función vectorial, $\mathbf{f}(\mathbf{x})$ es diferenciable podemos aproximarla en un punto $\mathbf{x}_0$ y en una dirección $\mathbf{v}$ por:

$$\mathbf{f}(\mathbf{x}_0+\mathbf{v}) \approx \mathbf{f}(\mathbf{x}_0) + \mathbf{Df}(\mathbf{x}_0)\,\mathbf{v}$$

donde $\mathbf{Df}(\mathbf{x}_0)$ es la matriz jacobiana:

$$\mathbf{Df}_{ij}(\mathbf{x}_0) = \frac{\partial f_i(\mathbf{x_0})}{\partial x_j}$$

Cuestión: Ya conoces con seguridad de Física al menos una matriz jacobiana. ¿Sabes cuál?

¡La velocidad!
Si tenemos una trayectoria de un móvil en el espacio en función del tiempo:

$$\mathbf{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

En la noticación anterior, la matriz jacobiana es:

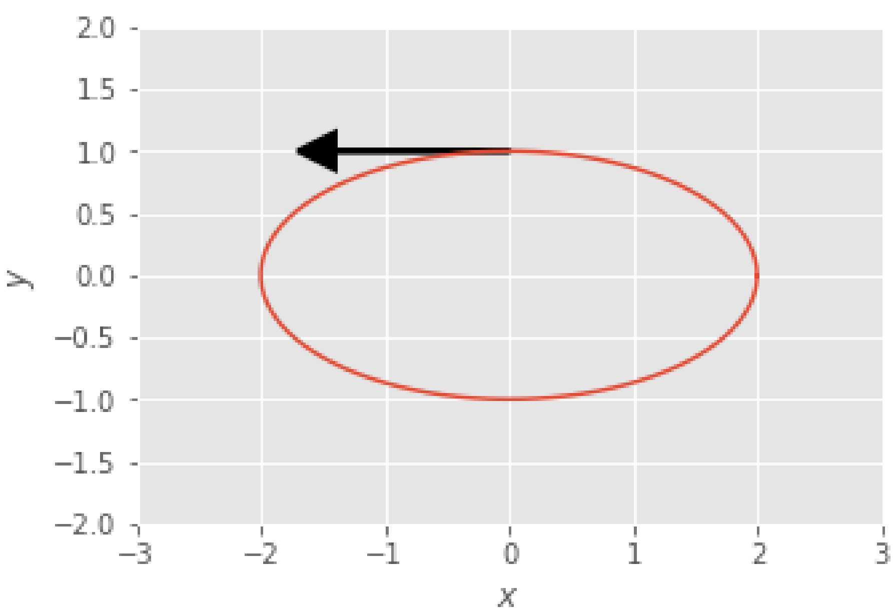
$$\mathbf{Dr}(t) = \mathbf{r}'(t) = \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}$$

El desarrollo de Taylor es por lo tanto:

$$\mathbf{r}(t_0 + \Delta t) \approx \mathbf{r}(t_0) + \mathbf{r}'(t_0)\Delta t$$

Ejemplo: La siguiente celda dibuja la línea de un movil a lo largo de una elipse y en un punto su velocidad (su "matriz" jacobiana)
Explora: puedes cambiar el valor de t_0 y observa como cambia la velocidad.

```
In [6]: t0 = np.pi/2.
        a, b, w = 2, 1, 1
        fx  = lambda t:    a*np.cos(w*t)
        fy  = lambda t:    b*np.sin(w*t)
        fxp = lambda t: -w*a*np.sin(w*t)
        fyp = lambda t:  w*b*np.cos(w*t)
        gf.line2d(fx, fy, length = 2.*np.pi);
        gf.arrow(fx(t0), fy(t0), fxp(t0), fyp(t0));
        plt.xlim(-a-1, a+1); plt.ylim(-b-1, b+1);
```



5.4 Función diferenciable

Hemos visto como son las funciones, diferenciable, pero ¿cómo las definimos matemáticamente? ¿cuándo una función será diferenciable?

Sea $\mathbf{f} : D \subset \mathbf{R}^n \to \mathbf{R}^m$, decimos que la función es **diferenciable** en un punto $\mathbf{x}$ interior a D si en una bola centrada en $\mathbf{x}$ de radio r y para los vectores que cumplen $\|\mathbf{v}\| < r$, existe una *transformación lineal* $\mathbf{T_x} : \mathbf{R}^n \to \mathbf{R}^m$ y una *función error*, $\mathbf{E} : \mathbf{R}^n \to \mathbf{R}^m$, $\mathbf{E}(\mathbf{x}, \mathbf{v})$, tales que:

$$\mathbf{f}(\mathbf{x} + \mathbf{v}) = \mathbf{f}(\mathbf{x}) + \mathbf{T_x}(\mathbf{v}) + \|\mathbf{v}\|\mathbf{E}(\mathbf{x}, \mathbf{v})$$

donde $\mathbf{E}(\mathbf{x}, \mathbf{v})$ es una función error de orden $\mathcal{O}(\|\mathbf{v}\|)$, que cumple:

$$\lim_{\mathbf{v} \to \mathbf{0}} \mathbf{E}(\mathbf{x}, \mathbf{v}) = \mathbf{0}$$

Esto es:

$$\lim_{\mathbf{v} \to \mathbf{0}} \frac{\mathbf{f}(\mathbf{x} + \mathbf{v}) - \mathbf{f}(\mathbf{x}) - \mathbf{T_x}(\mathbf{v})}{\|\mathbf{v}\|} = \lim_{\mathbf{v} \to \mathbf{0}} \mathbf{E}(\mathbf{x}, \mathbf{v}) = \mathbf{0}$$

A $\mathbf{T_x}$ se llama diferencial o **derivada** de $\mathbf{f}$ en $\mathbf{x}$.

La expresión anterior es el **desarrollo de Taylor** de primer orden de $\mathbf{f}(\mathbf{x} + \mathbf{v})$

$\mathbf{T_x}$ es la **derivada** de $\mathbf{f}$ en $\mathbf{x}$.

Si se trata de una función escalar, $f(\mathbf{x})$, es un **gradiente**, $\nabla f(\mathbf{x})$.

Si se trata de una función vectorial, $\mathbf{f}(\mathbf{x})$, es **la matriz jacobiana**, $\mathbf{Df}(\mathbf{x})$

Y te adelanto que:

$\mathbf{T_x}(\mathbf{v})$ es la **derivada direccional** de $\mathbf{f}(\mathbf{x})$ en el punto $\mathbf{x}$ y la dirección $\mathbf{v}$

Si se trata de un función escalar, $\mathbf{T_x}(\mathbf{v}) = \nabla f(\mathbf{x})\mathbf{v}$.

Si se trata de una función vectorial, $\mathbf{T_x}(\mathbf{v}) = \mathbf{Df}(\mathbf{x})\mathbf{v}$

Propiedades de las funciones diferenciables

Ejercicios: Intenta ahora demostrar los siguientes teoremas:

Teorema

Si una función $\mathbf{f}(\mathbf{x})$ de $\mathrm{R}^n \to \mathrm{R}^m$ es diferenciable en un punto interior $\mathbf{x}$, entonces en continua en ese punto.

Teorema

Si una función $\mathbf{f}(\mathbf{x})$ de $\mathrm{R}^n \to \mathrm{R}^m$ es diferenciable en un punto interior $\mathbf{x}$, su derivada direccional a lo largo de un vector $\mathbf{v}$ viene dada por:

si la función es escalar

$$f'(\mathbf{x}; \mathbf{v}) = \nabla f(\mathbf{x})\,\mathbf{v}$$

si la función es vectorial

$$\mathbf{f}'(\mathbf{x}; \mathbf{v}) = \mathbf{Df}(\mathbf{x})\,\mathbf{v}$$

Teorema: Sean $\mathbf{f}, \mathbf{g}$ dos funciones vectoriales definidas en $S \subset \mathbf{R}^n \to \mathbf{R}^m$; f, g dos campos escalares definidos en $S' \subset \mathbf{R}^n \to \mathbf{R}$ y λ un número real. Si son diferenciables en un punto $\mathbf{x}$, la siguientes funciones también lo son, con la siguiente derivada:

$$i) \quad \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}); \quad \mathbf{D}[\mathbf{f}+\mathbf{g}](\mathbf{x}) = \mathbf{Df}(\mathbf{x}) + \mathbf{Dg}(\mathbf{x})$$

$$ii) \quad \lambda\, \mathbf{f}(\mathbf{x}); \quad \mathbf{D}[\lambda\, \mathbf{f}](\mathbf{x}) = \lambda\, \mathbf{Df}(\mathbf{x})$$

$$iii) \quad f(\mathbf{x})\, g(\mathbf{x}); \quad \nabla[f(\mathbf{x})\, g(\mathbf{x})] = \nabla f(\mathbf{x})\, g(\mathbf{x}) + f(\mathbf{x})\, \nabla g(\mathbf{x})$$

$$iv) \quad f(\mathbf{x})/g(\mathbf{x}); \quad si\ g(\mathbf{x}) \neq 0 \quad \nabla[f(\mathbf{x})/g(\mathbf{x})] = \frac{\nabla f(\mathbf{x}) g(\mathbf{x}) - f(\mathbf{x}) \nabla g(\mathbf{x})}{g^2(\mathbf{x})}$$

Condición suficiente de diferenciabilidad Si recuerdas un función real de una dimensión, $f(x)$, era diferenciable, tenía desarrollo de Taylor, en un punto interior, x, si su derivada, $f'(x)$, era continua en un intervalo centrado en ese punto x.

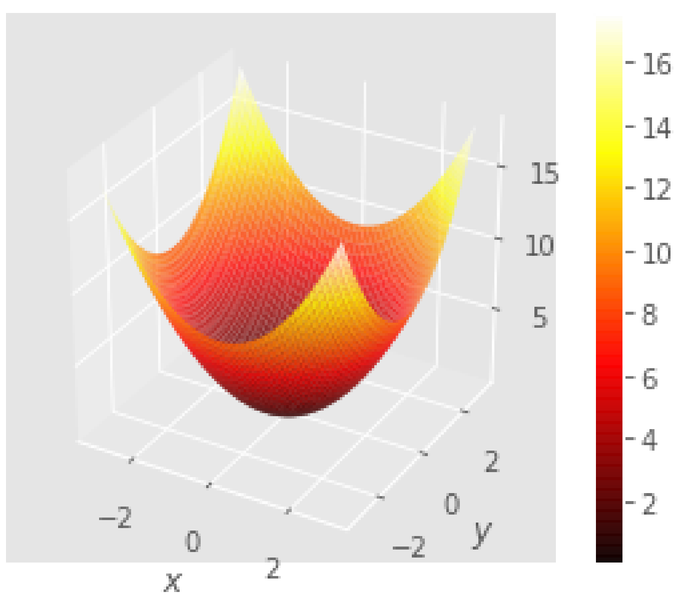
Ahora para funciones escalares, se cumple:

Teorema:

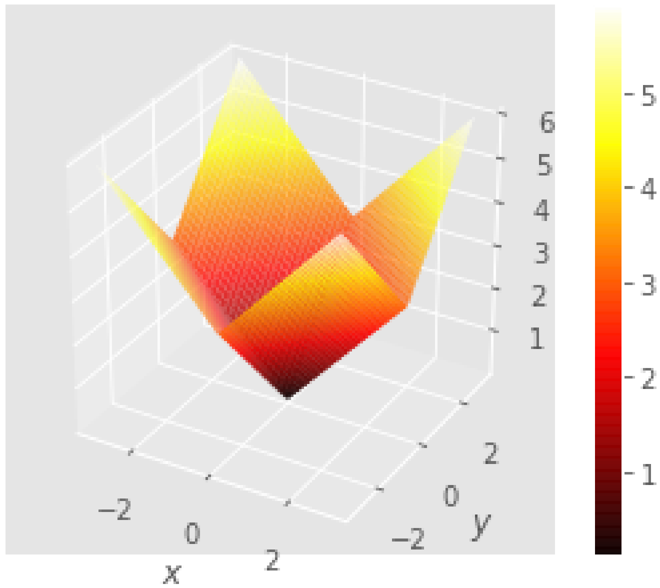
Una función $f(\mathbf{x})$ definida en un dominio $D \subset \mathrm{R}^n \to \mathrm{R}$ es continua si sus derivadas parciales, $f'(\mathbf{x})$, son continuas en una bola en torno a $\mathbf{x}$.

Cuestión: En la siguiente celda están dibujadas dos funciones escalares, $f(x, y)$ de $\mathrm{R}^2 \to \mathrm{R}$, una de ellas es diferenciable en todo el dominio y la otra no. ¿Sabes cual es la diferenciable?

```
In [7]: fun1 = lambda x, y : x*x + y*y
        gf.graph(fun1);
```



```
In [8]: fun2 = lambda x, y : np.abs(x) + np.abs(y)
        gf.graph(fun2);
```



¡Esto es todo por ahora!

5.5 Apendices

Demostración de la condición suficiente de diferenciabilidad

Lo demostraremos solamente en funciones escalares.

Queremos comprobar que

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) - \nabla f(\mathbf{x})\mathbf{v} = \mathbf{E}(\mathbf{x},\mathbf{v})$$

y que

$$\lim_{\mathbf{v}\to\mathbf{0}} \mathbf{E}(\mathbf{x},\mathbf{v}) = 0$$

Sea el vector $\mathbf{v}$ con norma λ y vector unitario $\mathbf{u}$. Construimos $n+1$ vectores $\mathbf{v}_i$, que van incorporando sucesivamente cada coordenada de $\mathbf{v}$, empezando por $\mathbf{v}_0 = \mathbf{0}$.

$$\mathbf{v} = \lambda\mathbf{u},\ \text{con}:\ \parallel \mathbf{u} \parallel = 1,\ \parallel \mathbf{v} \parallel = \lambda;\ \mathbf{v} = \lambda \sum_{k=1}^{n} u_k \mathbf{e}_k$$

$$\mathbf{v}_0 = \mathbf{0},\ \mathbf{v}_k = \sum_{i=1}^{k} \lambda u_k;\ \ k = 1,\ldots,n;\ \mathbf{v}_n = \mathbf{v}$$

Consideramos, por comodidad, los vectores:

$$\mathbf{b}_k = \mathbf{x} + \mathbf{v}_k,\ \mathbf{b}_k = \mathbf{b}_{k-1} + \lambda u_k \mathbf{e}_k,\ k = 1,\ldots,n$$

Reescribimos la parte izquierda de la igualdad con una suma telescópica:

$$f(\mathbf{x}+\mathbf{v})-f(\mathbf{x})=f(\mathbf{x}+\mathbf{v}_n)-f(\mathbf{x}+\mathbf{v}_{n-1})+\cdots+f(\mathbf{x}+\mathbf{v}_1)-f(\mathbf{x}+\mathbf{v}_0)$$

$$=\sum_{k=1}^{n} f(\mathbf{x}+\mathbf{v}_k)-f(\mathbf{x}+\mathbf{v}_{k-1})=\sum_{k=1}^{n} f(\mathbf{b}_k)-f(\mathbf{b}_{k-1})=\sum_{k=1}^{n} f(\mathbf{b}_{k-1}+\lambda u_k\mathbf{e}_k)-f(\mathbf{b}_{k-1})$$

Cada sumando del sumatorio, por ejemplo: $f(\mathbf{b}_{k-1}+\lambda u_k\mathbf{e}_k)-f(\mathbf{b}_{k-1})$, es simplemente una función de una dimensión, en la coordenada k. Podemos aplicar el teorema del valor medio, un punto $\mathbf{c}_k$ entre $\mathbf{b}_{k-1}+\lambda u_k\mathbf{e}_k$ y $\mathbf{b}_{k-1}$, y siendo la anchura del intervalo λu_k:

$$=\sum_{k=1}^{n} f(\mathbf{b}_{k-1}+\lambda u_k\mathbf{e}_k)-f(\mathbf{b}_{k-1})=\sum_{k=1}^{n}\lambda u_k f_k'(\mathbf{c}_k)$$

Ahora introducimos un sumando nulo, $\sum_{k=1}^{n}\lambda u_k\,[f_k'(\mathbf{x})-f_k'(\mathbf{x})]$, con el fin de obtener el término $\nabla f(\mathbf{x})\,\mathbf{v}$:

$$=\sum_{k=1}^{n}\lambda u_k\, f_k'(\mathbf{c}_k)+\sum_{k=1}^{n}\lambda u_k\,[f_k'(\mathbf{x})-f_k'(\mathbf{x})]$$

$$=\nabla f(\mathbf{x})\cdot\mathbf{v}+\lambda\sum_{k=1}^{n}[f_k'(\mathbf{c}_k)-f_k'(\mathbf{x})]u_k$$

Si recuperamos la expresión de función diferenciable:

$$f(\mathbf{x}+\mathbf{v})=f(\mathbf{x})+\nabla f(\mathbf{x})\,\mathbf{v}+\parallel\mathbf{v}\parallel E(\mathbf{x},\mathbf{v});\ \text{donde}: E(\mathbf{x},\mathbf{v})=\sum_{k=1}^{n}[f_k'(\mathbf{c}_k)-f_k'(\mathbf{x})]u_k$$

Cuando $\mathbf{v}\to 0$, $\lambda\to 0$, entonces $\forall k$, $\mathbf{c}_k\to\mathbf{x}$. *Si las derivadas parciales, $f_k'(\mathbf{x})$, son continuas* entoces $\mathbf{x}$, $E(\mathbf{x},\mathbf{v})\to 0$. Luego la función es diferenciable. Q.E.D