tderivadas_gradiante_jacobiana

February 8, 2019

1 Métodos III - Derivadas

1.1 Gradiente, matriz jacobina, 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  Fri Feb  8 10:22:01 2019
```

1.1.1 Objectivos

Presentar los conceptos de

- gradiente y matriz jacobiana
- definición de función diferenciable

Discutir:

- Condición suficiente de diferenciabilidad
- Bondades de las funciones diferenciables

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

        # numpy and matplotlib
        import numpy as np
        import matplotlib
        import matplotlib.pyplot as plt

        from mpl_toolkits.mplot3d import Axes3D
        # possible styles: ggplot (simplicity), bmh (scientify data),
        matplotlib.style.use('ggplot')
```

1.1.2 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:

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

Donde la forma:

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

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

$$\nabla f(x_0,y_0) = \left(f'_x(x_0,y_0),\, f'_y(x_0,y_0)\right)$$

y que asociaremos con el símbolo nabla, ∇.

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

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:

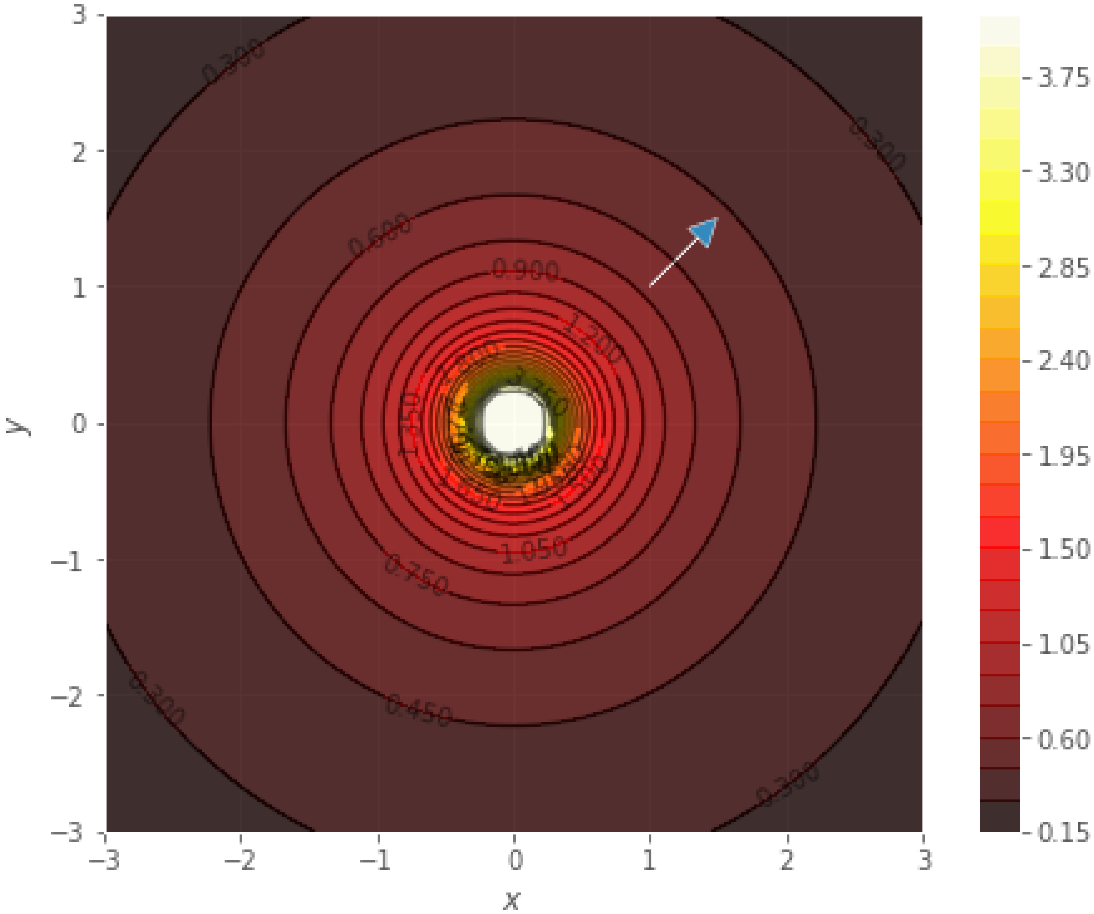
La siguiente figura nos muestra los conjuntos de nivel del potencial eléctrico de una carga. También nos muestra el campo eléctrico, (el gradiente cambiado de signo), en un determinado punto (x_0, y_0)

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

```
In [3]: def gradient_charge_potential(x0 = 2., y0 = 2., V_cutoff = 4.):
            fig, ax = plt.subplots(figsize=(8, 6))
            xs = np.linspace(-3., 3., 100)
            ys = np.linspace(-3., 3., 100)
            xms, yms = np.meshgrid(xs, ys)
            zms = 1./np.sqrt(xms*xms+yms*yms)
            sel = zms > V_cutoff; zms[sel] = V_cutoff
            r0 = np.sqrt(x0*x0+y0*y0)
            vx, vy = x0/(r0**3), y0/(r0**3)
            # color fill contour
            c0 = ax.contourf(xms, yms, zms, 30, alpha=0.8, cmap='hot')
            ax.arrow(x0, y0, vx, vy, head_width=0.2, head_length=0.2,)
            c1 = ax.contour(xms, yms, zms, 30, colors='black', alpha=0.5);
```

```
    c1.clabel(fontsize=10, inline=1)
    ax.set_xlabel(r'$x$'); ax.set_ylabel(r'$y$'); ax.set_aspect('equal')
    fig.colorbar(c0, ax=ax);
    return
```

```
In [4]: x0, y0 = 1., 1.
        gradient_charge_potential(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, luego el término es máximo.

O lo que es lo mismo, si el desplazamiento $\mathbf{v}$ 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 si nos movemos 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)

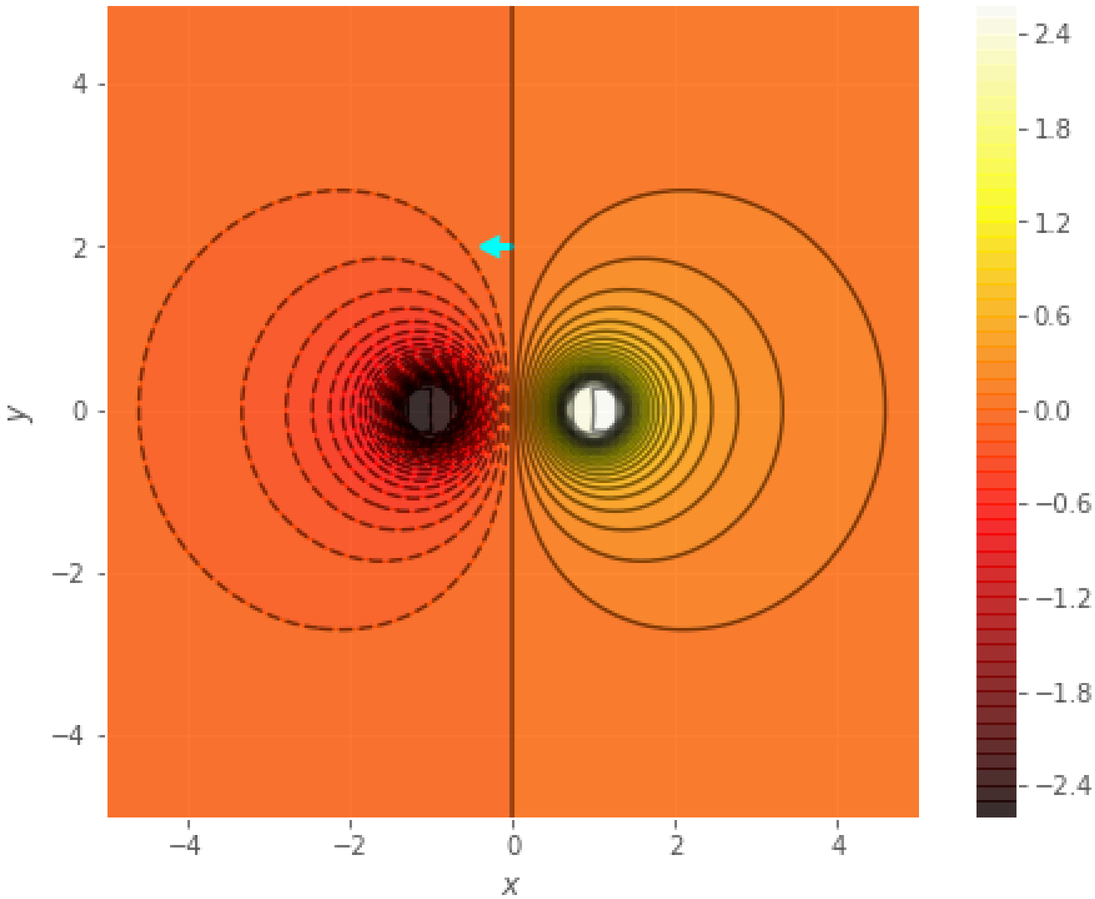
Experimenta y cambia el punto y observa como cambia el gradiente.

El gradiente debe indicarnos la dirección en la que cambía más fuertemente el potencia eléctrico, y también debe ser normal a los conjuntos de nivel, a las líneas equipotenciales.

```
In [5]: def gradient_dipole_potential(x0, y0, d = 2., V_cutoff = 3.):
            qpos, xpos, ypos = +1,  d/2, 0.
            qneg, xneg, yneg = -1, -d/2, 0.
            def V(x, y, q, xi, yi):
                dx, dy = (x-xi), (y-yi)
                ri      = np.sqrt(dx*dx + dy*dy)
                vi      = 1/ri
                vi[ vi > V_cutoff] = V_cutoff
                return  q*vi
            def E(x, y, q, xi, yi):
                dx, dy = (x-xi), (y-yi)
                ri      = np.sqrt(dx*dx + dy*dy)
                Ex, Ey = q*dx/(ri*ri*ri), q*dy/(ri*ri*ri)
                return np.array([Ex, Ey])
            xs = np.linspace(-5., 5., 100)
            ys = np.linspace(-5., 5., 100)
            xms, yms = np.meshgrid(xs, ys)
            Vdipole = V(xms, yms, qpos, xpos, ypos) + V(xms, yms, qneg, xneg, yneg)
            Edipole = E(x0 , y0, qpos, xpos, ypos)   + E(x0 , y0, qneg, xneg, yneg)

            vx, vy = Edipole[0], Edipole[1]
            fig, ax = plt.subplots(figsize=(8, 6))
            c0 = ax.contourf(xms, yms, Vdipole, 60, alpha=0.8, cmap='hot')
            c1 = ax.contour(xms, yms, Vdipole, 60, colors='black', alpha=0.5);
            ax.arrow(x0, y0, vx, vy, head_width=0.2, head_length=0.2, lw= 2, color='cyan')
            #c1.clabel(fontsize=10, inline=1)
            ax.set_xlabel(r'$x$'); ax.set_ylabel(r'$y$'); ax.set_aspect('equal')
            fig.colorbar(c0, ax=ax);
            return

In [6]: x0, y0 = 0, 2.
        gradient_dipole_potential(x0, y0)
```



1.1.3 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 $R^n \to R^m$, que es diferenciable en un punto $\mathbf{x}_0$

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

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

O también:

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

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?
ąLa 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:

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

El desarrollo de Taylor es por lo tanto:

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

1.1.4 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* $\mathbf{T}_\mathbf{x} : \mathbf{R}^n \to \mathbf{R}^m$ y una función, $\mathbf{E} : \mathbf{R}^n \to \mathbf{R}^m$, $\mathbf{E}(\mathbf{x}, \mathbf{v})$, tales que:

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

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

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

Esto es:

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

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

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

$\mathbf{T}_\mathbf{x}$ 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:

$\mathbf{T}_\mathbf{x}(\mathbf{v})$ 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, $\mathbf{T}_\mathbf{x}(\mathbf{v}) = \nabla f(\mathbf{x})\mathbf{v}$.

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

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

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

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 $\mathbf{x}$.

Y para las funciones vectoriales:

Teorema:

Una función vectorial $\mathbf{f}(\mathbf{x})$ definida en un dominio $D \subset \mathrm{R}^n \to \mathrm{R}^m$, es diferenciable, si y solo si, lo son cada una de sus funciones componentes, $f_i(\mathbf{x}),\ i = 1, \ldots, m$.

ąEsto es todo por ahora!