Tarea 5 Optimización

Roberto Vásquez Martínez
Profesor: Joaquín Peña Acevedo

13/Marzo/2022

1 Ejercicio 1 (6 puntos)

Programar el método de descenso máximo con tamaño de paso fijo y probarlo.

El algoritmo recibe como parámetros la función gradiente $g(x)$ de la función objetivo, un punto inicial x_0, el valor del tamaño de paso α, un número máximo de iteraciones N, la tolerancia $\tau > 0$. Fijar $k = 0$ y repetir los siguientes pasos:

1. Calcular el gradiente g_k en el punto x_k, $g_k = g(x_k)$.
2. Si $\|g_k\| < \tau$, hacer $res = 1$ y terminar.
3. Elegir la dirección de descenso como $p_k = -g_k$.
4. Calcular el siguiente punto de la secuencia como

$$x_{k+1} = x_k + \alpha p_k$$

5. Si $k + 1 \geq N$, hacer $res = 0$ y terminar.
6. Si no, hacer $k = k + 1$ y volver el paso 1.
7. Devolver el punto x_k, g_k, k y res.

De acuerdo con la proposición vista en la clase 12, para que el método de máximo descenso con paso fijo para funciones cuadráticas converja se requiere que el tamaño de paso α cumpla con

$$0 < \alpha < \frac{2}{\lambda_{\max}(A)} = \alpha_{\max},$$

donde $\lambda_{\max}(A)$ es el eigenvalor más grande de A.

1. Escriba una función que implementa el algoritmo de descenso máximo con paso fijo.

2. Programe las funciones cuadráticas y sus gradientes

$$f_i(x) = \frac{1}{2}x^\top \mathbf{A}_i x - \mathbf{b}_i^\top x, \quad i = 1, 2$$

   donde

$$\mathbf{A}_1 = \begin{bmatrix} 1.18 & 0.69 \\ 0.69 & 3.01 \end{bmatrix}, \qquad \mathbf{b}_1 = \begin{pmatrix} -0.24 \\ 0.99 \end{pmatrix}.$$

   y

$$\mathbf{A}_2 = \begin{pmatrix} 6.36 & -3.07 & -2.8 & -3.42 & -0.68 \\ -3.07 & 10.19 & 0.74 & 0.5 & 0.72 \\ -2.8 & 0.74 & 4.97 & -1.48 & 1.93 \\ -3.42 & 0.5 & -1.48 & 4.9 & -0.97 \\ -0.68 & 0.72 & 1.93 & -0.97 & 3.21 \end{pmatrix}, \qquad \mathbf{b}_2 = \begin{pmatrix} 0.66 \\ 0.37 \\ -2.06 \\ 0.14 \\ 1.36 \end{pmatrix}.$$

3. Fije el número máximo de iteraciones $N = 2000$ y la tolerancia $\tau = \sqrt{\epsilon_m}$, donde ϵ_m es el épsilon de la máquina. Para cada función cuadrática, calcule $\alpha_{\max}$ de la matriz $\mathbf{A}_i$. Pruebe con los tamaños de paso α igual a $1.1\alpha_{\max}$ y $0.9\alpha_{\max}$. Use el punto inicial

$$\mathbf{x}_0 = \begin{pmatrix} -38.12 \\ -55.87 \end{pmatrix} \quad \text{para} \quad f_1$$

$$\mathbf{x}_0 = \begin{pmatrix} 4.60 \\ 6.85 \\ 4.31 \\ 4.79 \\ 8.38 \end{pmatrix} \quad \text{para} \quad f_2$$

4. En cada caso imprima x_k, $\|g_k\|$, el número de iteraciones k y el valor de *res*.

1.1 Solución:

El número de iteraciones que consideraremos en el método de descenso máximo con paso fijo será $N = 2000$ y la tolerancia $\tau = \sqrt{\epsilon_m}$, donde ϵ_m es el épsilon de la máquina.

Calcularemos $\alpha_{\max}$ para cada función cuadrática f_i. Probaremos el método de descenso máximo con paso fijo con los tamaños de paso $0.9\alpha_{\max}$ y $1.1\alpha_{\max}$.

1.1.1 Función cuadrática 1

Haremos las pruebas correspondientes a la función f_1. En primer lugar, determinamos $\alpha_{\max}$, además de validar si $\mathbf{A}_1$ es positiva definida.

[75]:
```python
import lib_t5
import importlib
importlib.reload(lib_t5)
from lib_t5 import *

# Iteraciones maximas, tolerancia
N=2000
tol=np.finfo(float).eps**(1/2)

# Matriz y vector modelo cuadrático
A1,b1=np.array([[1.18,0.69],[0.69,3.01]]),np.array([-0.24,0.99]).reshape(-1,1)
# Condición inicial
x0=np.array([-38.12,-55.87]).reshape(-1,1)
```

```
# Verificar si es positiva definida
eig_val_A1=np.real_if_close(np.linalg.eigvals(A1))
print('El valor eigenvalor mínimo es: ',np.min(eig_val_A1))
# Calcula a_max
a_max_1=2.0/np.max(eig_val_A1)
print('a_max = ',a_max_1)
```

```
El valor eigenvalor mínimo es:  0.9489960733051563
a_max =  0.6170927420133021
```

De lo anterior tenemos que $\mathbf{A}_1$ es positiva definida, luego el punto crítico es mínimo global. Ahora haremos las pruebas con los distintos tamaños de pasos.

$0.9\alpha_{\max}$ Para este tamaño de paso, imprimimos el número de iteraciones, el valor en el mínimo y la distancia respecto al punto crítico para validar que en efecto convergemos al mínimo global

[76]:
```
proof_grad_max_fix_quad(quad_fun,grad_quad_fun,x0,N,tol,0.9*a_max_1,args=[A1,b1])
```

```
res =  1
El método de descenso máximo con paso fijo CONVERGE
k =  105
fk =  -0.26949669993822545
||gk|| =  1.4137071703555776e-08
xk =  [-0.45696914  0.43365738]
||xk-x*|| =  4.361942126279292e-09
```

La convergencia esta asegurada por la Proposición de la Clase 12 para cualquier condición inicial pues $0.9\alpha_{\max} < \alpha_{\max}$, y en este caso la convergencia fue relativamente rápida.

$1.1\alpha_{\max}$ En este caso no tenemos segura la convergencia, realizamos la misma prueba para este tamaño de paso

[78]:
```
proof_grad_max_fix_quad(quad_fun,grad_quad_fun,x0,N,tol,1.1*a_max_1,args=[A1,b1])
```

```
res =  0
El método de descenso máximo con paso exacto NO CONVERGE
k =  2000
fk =  inf
||gk|| =  inf
xk =  [-4.77996787e+159 -1.42775834e+160]
||xk-x*|| =  inf
```

En este caso por completo divergemos, el valor de la función aumenta cada vez más, incluso diverge la norma del gradiente. Nos alejamos tanto del punto crítico que la distancia final respecto al punto crítico es ∞.

El cambio fue sutil en el tamaño de paso fijo, pero nos llevó a un resultado degenerado.

1.1.2 Función cuadrática 2

Al igual que la función f_1, primero validamos si $\mathbf{A}_2$ es positiva definida además de calcular $\alpha_{\max}$ para esta función cuadrática.

[79]:
```
# Matriz y vector modelo cuadrático
A2,b2=np.array([[6.36,-3.07,-2.8,-3.42,-0.68],[-3.07,10.19,0.74,0.5,0.72],[-2.
↪8,0.74,4.97,-1.48,1.93],[-3.42,0.5,-1.48,4.9,-0.97],[-0.68,0.72,1.93,-0.97,3.
↪21]]),np.array([0.66,0.37,-2.06,0.14,1.36]).reshape(-1,1)

# Condición inicial
x0=np.array([4.60,6.85,4.31,4.79,8.38]).reshape(-1,1)

# Verificar si es positiva definida
eig_val_A2=np.real_if_close(np.linalg.eigvals(A2))
print('El valor eigenvalor mínimo es: ',np.min(eig_val_A2))

# Calcula a_max
a_max_2=2.0/np.max(eig_val_A2)
print('a_max = ',a_max_2)
```

```
El valor eigenvalor mínimo es:  0.12547412774810465
a_max =  0.1529725843120685
```

Del resultado anterior concluimos que $\mathbf{A}_2$ es una matriz simétrica positiva definida, luego el punto crítico es óptimo global. Haremos las pruebas con los distintos tamaños de pasos.

$0.9\alpha_{\max}$ Imprimimos el número de iteraciones, el valor en el mínimo y la distancia respecto al punto crítico para validar que en efecto convergemos al mínimo global.

[80]:
```
proof_grad_max_fix_quad(quad_fun,grad_quad_fun,x0,N,tol,0.9*a_max_2,args=[A2,b2])
```

```
res =  1
El método de descenso máximo con paso fijo CONVERGE
k =  1064
fk =  -2.6497175235052906
||gk|| =  1.4711921029622859e-08
xk =  [-2.77194407 -0.52190805 -3.05959477 -2.57614049  1.01464594]
||xk-x*|| =  1.1725061853105777e-07
```

Ya sabíamos que la convergencia se obtenía por la proposición de la Clase 12, en este caso tomó más iteraciones que con la función f_1 pero esto se puede deber a que $\alpha_{\max}$ para f_2 es más pequeño que el correspondiente a f_1.

$1.1\alpha_{\max}$ En este caso no tenemos segura la convergencia, realizamos la misma prueba para este tamaño de paso

[81]:
```
proof_grad_max_fix_quad(quad_fun,grad_quad_fun,x0,N,tol,1.1*a_max_2,args=[A2,b2])
```

```
res =  0
El método de descenso máximo con paso exacto NO CONVERGE
k =  2000
fk =  inf
||gk|| =  inf
xk =  [-7.41437598e+158  9.63088421e+158  3.28353818e+158  2.91033121e+158
  1.57034121e+158]
||xk-x*|| =  inf
```

Al igual que con la función f_1, con el tamaño de paso $1.1\alpha_{\max}$ el método de descenso máximo con paso fijo diverge, tanto que la norma del gradiente y el valor de la función en la última iteración es infinito, lo que puede significar que con este tamaño de paso, en lugar de reducir el valor de la función este se incrementa cada vez más, a pesar de estar considerando direcciones de descenso pues seguro no se satisface la condición de descenso suficiente.

1.2 Ejercicio 2 (4 puntos)

Pruebe el método de descenso máximo con paso fijo aplicado a la función de Rosenbrock.

Encuentre un valor adecuado para α para que el algoritmo converja. Use como punto inicial el punto $(-12, 10)$.

Imprima x_k, $\|g_k\|$, el número de iteraciones k y el valor de *res*.

1.2.1 Solución:

Haremos la prueba del método de descenso máximo con paso fijo con la función de Rosenbrock, del que sabemos tiene su mínimo en $x_* = (1, 1)$

[82]:
```
importlib.reload(lib_t5)
from lib_t5 import *
x0=np.array([-12.0,10.0]).reshape(-1,1)
N=1e7
proof_grad_max_fix(f_Rosenbrock,grad_Rosenbrock,x0,N,tol,a=0.000036,args=None)
```

```
res =  1
El método de descenso máximo con paso fijo CONVERGE
k =  1407272
fk =  2.77999204102933e-16
||gk|| =  1.4901139451473246e-08
xk =  [0.99999998 0.99999997]
```

En efecto, hemos alcanzado convergencia en $k = 1,407,272$ iteraciones con tamaño de paso fijo $\alpha = 3.5 \times 10^{-5}$. Como observamos en la tarea anterior, el punto inicial $x_0 = (-12, 10)$ se encuentra en un valle y relativamente alejado del óptimo, incluso con backtracking no pudimos hallar un tamaño de paso eficiente, así que recurrimos a un tamaño de paso pequeño pero aumentando el número de iteraciones máximas para lograr la convergencia. El tiempo de cómputo para la convergencia fue de 25.4s.

Cabe resaltar que con tamaños de paso un poco más grandes se obtienen `NaN`'s al momento de correr el algoritmo, por eso un tamaño de paso chico fue una mejor opción.