



ESCUELA SUPERIOR DE FÍSICA Y MATEMÁTICAS

INSTITUTO POLITÉCNICO NACIONAL

Tarea 7 - Optimización No Lineal

CAMARA MEDINA CYNTHIA LILIAN
REYES ZAMORA OLLIN
ALANIS GONZÁLEZ EDZON OMAR
MENDOZA URRUSQUIETA JAIR NATAEL

20 de abril de 2023

Preguntas

1. **¿Por qué es válido analizar la convergencia de los métodos vistos en clase solo para funciones cuadráticas?**

Debido a que el método de Newton es exacto cuando el orden de las funciones de 2, entonces la primera aproximación es exacta, y el método produce el vector mínimo verdadero en un paso. Debido a que la función de aproximación tiene forma cuadrática.

2. **Enumere las posibles desventajas del método de Newton para resolver problemas de optimización con varias variables.**

- I. La convergencia del método de Newton suele ser en razón a la cercanía de punto inicial al óptimo de la función.
- II. El costo computacional que requiere es bastante alto debido a que necesita información de la primera y segunda derivada, por lo que en funciones de alta complejidad este tiende a ser lento o divergente.
- III. Si la Hessiana de la función no es definida positiva no hay garantía de que el método de Newton produzca en todos los casos una sucesión de valores decrecientes de la función objetivo.
- IV. La resolución de un sistema de ecuaciones en cada iteración es muy costosa computacionalmente.

3. **Explica cuál es la razón de que existan modificaciones al método de Newton para optimización.**

Debido a que el método de Newton original necesita información con un gran costo computacional, se requiere reducir este costo de tal manera que este sea mejor en comparación al original. Además de asegurar que el “paso” que se dé sea de descenso debido a que el método no asegura que esto sea verás en cada iteración.

4. **¿Cuál es la importancia del método de Gradientes Conjugados respecto a los métodos estudiados anteriormente?**

El método de gradientes conjugados representa el avance que se ha dado en relación con los métodos visto anteriormente, debido a que su costo computacional, si eficiencia, eficacia y robustez, es mejor que algunos, en algunos ámbitos, pero en relación a todas estas características es uno de los mejores vistos hasta ahora. Debido a que los métodos de direcciones conjugadas aprovechan información geométrica de la función para dar pasos idóneos considerando la concavidad de f . El algoritmo de gradientes conjugados es en sí mismo un método de direcciones conjugadas, y por lo tanto es capaz de minimizar una función cuadrática (cuya forma está dada en la matriz Q , en principio, definida positiva) de n variables en n pasos.

5. **Explique la filosofía que siguen los métodos de tipo Cuasi-Newton.**

Al igual que en el caso de Newton para una variable, donde el método de la secante busca aproximar la derivada de manera iterativa, evitando así calcularla numéricamen-

te, los métodos cuasi Newton buscan aproximar la matriz H , mediante matrices H_k utilizadas de manera iterativa.

Método de gradientes conjugados

Buscar el óptimo del siguiente problema mediante el método de gradientes conjugados, con todas las variantes para el cálculo de β vistas en clase.

Minimizar

$$f(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

Hacer cuatro iteraciones paso a paso, es decir, reportando a cada iteración k los valores de $x^{(k)}$, $f(x^{(k)})$, y $\nabla f(x^{(k)})$. Considere solo 4 cifras de precisión. Comience las iteraciones en el punto $x^{(0)} = (10, -10, 10, -10)$. Escriba sus resultados y conclusiones.

Fletcher-Reeves

$$g^k = \begin{bmatrix} 2(x_1 + 10x_2) + 40(x_1 - x_4)^3 \\ 20(x_1 + 10x_2) + 4(x_2 - 2x_3)^3 \\ 10(x_3 - x_4) - 8(x_2 - 2x_3)^3 \\ -10(x_3 - x_4) - 40(x_1 - x_4)^3 \end{bmatrix} \quad x^0 = [10 \quad -10 \quad 10 \quad -10]$$

k=0

$$\begin{aligned} x^1 &= [-2.1074 \quad -5.8433 \quad 1.8153 \quad 2.1218] \\ f(x^1) &= 14920.6991 \\ g^1 &= [-3146.84 \quad -4612.12 \quad 6799.56 \quad 3028.83] \end{aligned}$$

k=1

$$\begin{aligned} x^2 &= [0.4598 \quad -1.9202 \quad -3.9834 \quad -0.3457] \\ f(x^2) &= 1758.38 \\ g^2 &= [-16.5791 \quad 509.444 \quad -1804.95 \quad 15.4717] \end{aligned}$$

k=2

$$\begin{aligned} x^3 &= [-0.0932 \quad -2.6715 \quad -2.90 \quad 0.1207] \\ f(x^3) &= 860.119 \\ g^3 &= [-54.0079 \quad -413.683 \quad -275.169 \quad 30.5985] \end{aligned}$$

$$\boxed{k=3}$$

$$\begin{aligned}x^4 &= [0.3154 \quad 0.00616 \quad -1.2623 \quad -0.1282] \\f(x^4) &= 52.4207 \\g^4 &= [5.3544 \quad 87.8185 \quad -149.72 \quad 7.8493]\end{aligned}$$

Pollak-Ribiere

$$\begin{aligned}x^0 &= [10 \quad -10 \quad 10 \quad -10] \quad f(x^0) = 21420.100 \\ \nabla f(x^0) &= [319820 \quad -109800 \quad 216200 \quad -320200]^T\end{aligned}$$

$$\boxed{k=0}$$

$$\begin{aligned}x^1 &= [-2.1074 \quad -5.8433 \quad 1.8153 \quad 2.1218] \\f(x^1) &= 14920 \\g^1 &= [-3146.84 \quad -4612.12 \quad 6799.56 \quad 5028.83]\end{aligned}$$

$$\boxed{k=1}$$

$$\begin{aligned}x^2 &= [-0.2820 \quad -3.0544 \quad -2.3068 \quad 0.3673] \\f(x^2) &= 993.695 \\g^2 &= [-72.6167 \quad -601.358 \quad -57.0686 \quad 37.7087]\end{aligned}$$

$$\boxed{k=2}$$

$$\begin{aligned}x^3 &= [-0.2881 \quad -2.7646 \quad -2.1579 \quad 0.3928] \\f(x^3) &= 820.784 \\g^3 &= [-68.449 \quad -543.752 \quad -55.3672 \quad 38.1343]\end{aligned}$$

$$\boxed{k=3}$$

$$\begin{aligned}x^3 &= [-0.0038 \quad -0.7012 \quad -2.0162 \quad 0.2218] \\f(x^3) &= 197.432 \\g^3 &= [-14.4909 \quad 7.5478 \quad -318.108 \quad 22.8393]\end{aligned}$$

Hestenes-Stiefel

$$x^0 = [10 \quad -10 \quad 10 \quad -10] \quad f(x^0) = 21420.100$$
$$\nabla f(x^0) = [319820 \quad -109800 \quad 216200 \quad -320200]^T$$

k=0

$$x^1 = [-2.1074 \quad -5.8433 \quad 1.8153 \quad 2.1218]$$
$$f(x^1) = 14920$$
$$g^1 = [-3146.84 \quad -4612.12 \quad 6799.56 \quad 5028.83]$$

k=1

$$x^2 = [-0.2768 \quad -3.0562 \quad -2.3033 \quad 0.3621]$$
$$f(x^2) = 756.629$$
$$g^2 = [-57.0514 \quad -687.525 \quad 446.574 \quad 6.5163]$$

k=2

$$x^3 = [-0.8567 \quad -1.7767 \quad -2.1440 \quad 1.0935]$$
$$f(x^3) = 1073.34$$
$$g^3 = [-341.282 \quad -824.616 \quad 758.456 \quad 295.554]$$

k=3

$$x^4 = [-0.9012 \quad 1.5996 \quad -2.1650 \quad 1.1564]$$
$$f(x^4) = 575.496$$
$$g^4 = [-382.256 \quad -256.523 \quad -196.057 \quad 381.666]$$

Podemos concluir que el mejor método es el de Fletcher-Reeves por su rápida convergencia. En todos se cumple la propiedad de descenso excepto en Hestenes-Stiefel.

Comprobación de Métodos

No olvide adicionar al número total de llamadas a la función objetivo, las correspondientes al cálculo de las derivadas con Diferencias Finitas, en el método que aplique.

Gradientes Conjugados

1. Programar en MATLAB el método de gradientes conjugados, con la variante que usted decida de entre las que vimos en clase, para aproximar el mínimo del siguiente problema:

$$f(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5x_1}{\pi} - 6 \right)^2 + 10 \left(1 - \frac{1}{8\pi} \right) \cos(x_1) + 10$$

2. Reporte el resultado y compárelo con el obtenido al usar los métodos de la tarea anterior. La comparación es en términos del resultado obtenido y el número de evaluaciones de la función objetivo. Utilice el método de Diferencias Finitas para aproximar todas las derivadas. Detenga las iteraciones en todos los casos cuando la norma del gradiente sea menor que 10^{-9} . Inicie las iteraciones en el punto $x^{(0)} = (-4, 1)$.

La siguiente tabla muestra los resultados y la comparación entre ellos.

	Punto Óptimo x^*	Llamadas a la función	Notas
Gradientes Conjugados - Fletcher Reeves	[0.00107, 6.00021]	1215	
Newton - Raphson	[-3.14070, 7.27430]	120	Nuestra hipótesis es que se llega a un punto distinto debido a la distancia entre el punto inicial y el óptimo, pues esta es de 6.40, la que consideramos muy grande para aplicar el método.
Pendiente Máxima	[0.00049, 5.99997]	259	Se modificó el error máximo permitido a 1E-3 debido a que para un error de orden menor a este los valores resultantes eran demasiado pequeños, con lo que el algoritmo no funcionaba adecuadamente.
Cuasi-Newton Rango 1	[-3.352E-5, 5.99994]	260	

3. Escriba sus conclusiones.

El mejor método para este problema es Cuasi-Newton de Rango 1. Concluimos esto debido a que, poniendo como tolerancia un número más pequeño que en el método de Descenso con Pendiente Máxima sólo se realizó una llamada a la función más, además de obtener una mucho mejor aproximación al punto óptimo real.

Métodos cuasi-Newton

1. Programar en **MATLAB** los métodos cuasi-Newton (rango uno, DFP y BFGS), que vimos en clase, para aproximar el mínimo del siguiente problema:

$$f(x_1, x_2, x_3, x_4) = x_1^2 + 2x_2^2 + 3x_3^2 + 4x_4^2 + (x_1 + x_2 + x_3 + x_4)^4$$

Detenga las iteraciones en todos los casos cuando la norma del gradiente sea menor que 10^{-9} . Inicie las iteraciones en el punto $x^{(0)} = (10, 10, 10, 10)$. Para los métodos cuasi-Newton reinicie el algoritmo cada 4 iteraciones, de ser necesario.

El óptimo por **Rango 1** encontrado es

$$x^8 = 1 \times 10^{-4} \cdot [0.1374 \quad -0.1111 \quad -0.0153 \quad -0.0133]$$

con 8 iteraciones.

Para **BFGS**

$$x^8 = 1 \times 10^{-4} \cdot [0.1374 \quad -0.1111 \quad -0.0153 \quad -0.0133]$$

el mismo punto en las mismas iteraciones.

Y para DFP

$$x^8 = 1 \times 10^{-4} \cdot [0.1374 \quad -0.1111 \quad -0.0153 \quad -0.0133]$$

el mismo punto de nuevo con las mismas iteraciones, en este caso no hubo mayor convergencia al cambiar el rango.

2. Compare el resultado anterior para aplicar al mismo problema el método de Newton.
No es posible aplicar newton ya que nos encontramos con una matriz singular, y esto causa un problema al calcular la inversa ya que no existe, y el código no puede ser completado.
3. Compare de nuevo el desempeño de sus programas para cuasi-Newton y Newton con la siguiente función, con $n = 10$, y algún punto de inicio.

$$f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$

Tomando el punto de inicio

$$x^0 = [1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1]$$

Los métodos con el mismo error antes requerido, convergieron al mismo punto

$$x^0 = [1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1]$$

en

- Rango 1 \rightarrow 109 iteraciones
- BFGS \rightarrow 103 iteraciones
- DFP \rightarrow 100 iteraciones
- Newton \rightarrow converge a un punto distinto al que nunca llega, esto por la distancia del punto inicial.

4. Escriba sus conclusiones, en particular respecto a la elección del punto de inicio.

El punto de inicio significa mucho para el punto a converger, en el caso de los métodos de Cuasi-Newton no hubo ningún problema al llegar todos al mismo punto, y sus variaciones solo cambiaron las iteraciones en la que lo lograron, pero en el método de newton si afecta bastante la distancia del punto inicial, lo cual lo llevo a no converger por alternar entre un punto que no era el mínimo.