Tarea 9 - Optimización

Erick Salvador Alvarez Valencia

CIMAT A.C.,
erick.alvarez@cimat.mx

Resumen En el presente reporte se hablará sobre la implementación del algoritmo BFGS tanto con una versión con la matríz Hessiana analítica, como una versión con una aproximación de la misma pero usando el método de diferencias finitas. Se incluirá iba descripción de la implementación asi como los resultados de algunas pruebas realizadas con la función de Rosenbrock en dos y cien dimensiones.
Finalmente se darán algunas conclusiones.

Keywords: Método BFGS, Optimización de funciones, Diferencias finitas.

1. Método BFGS

Primeramente se aplicó el método BFGS el cual ess usado para la minimización de funciones. Dicho metodo era con la inversa de la matriz Hessiana, la cual esta asegurada que existe ya que esta es simétrica y definida positiva. La idea del método BFGS es trabajar con un modelo cuadratico que aproximada localmente a la función y por lo tanto usa la actualización de Newton para encontrar la siguiente dirección $p_k = -B_k^{-1}$ y como a mencionó anteriormente este método trabaja con la inversa de esta matriz para simplemente realizar el producto mostrado anteriormente y no tener que resolver el sistema de ecuaciones asociado. Para evitar que el proceso sea costoso comoutacionalmente en el sentido de calcular esta matriz en cada iteración del algoritmo se tiene una regla de actualización que solamente esta definida en términos de productos puntos y productos de matrices, la cual es la siguiente:

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \tag{1}$$

Donde $/rho_k$ esta definida como $\frac{1}{y_k^T s_k}$

Lo anterior se llama actualización del método BFGS.
Normalmente para calcular el siguiente valor de x_k se requiere de un tamaño de paso, el cual se calcula con el método de backtracking, esto para buscar un alpha que asegure suficiente descenso. Y para ello se incluyen las condiciones de Armijo o de Wolfe (para la implementación que se hizo en este trabajo se usaron las condiciones de Armijo).

2. Ejecuciones y resultados

Para este trabajo se hizo la implementación del metodo BFGS y se hicieron dos tipos de pruebas, la primera fue usando una matriz Hessiana analítico, y la segunda fue usando el enfoque de diferencias finitas para aproximar la Hessiana en un punto x_0. Dicho cálculo se hace de la siguiente manera:

$$h_{ij} = \frac{f(x + he_i + he_j) - f(x + he_i) - f(x + he_j) + f(x)}{h^2} \tag{2}$$

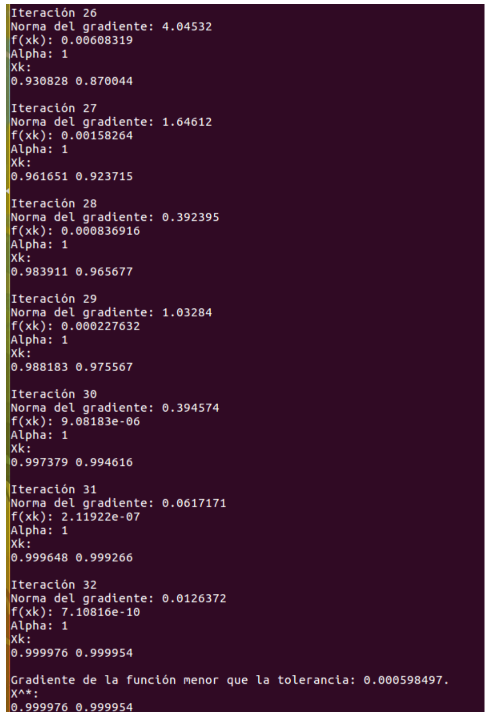
En cada una de esas pruebas se fue capturando el numero de iteracion, la norma del gradiente, el valor de la función, el valor del tamaño de paso y el valor de x, todo lo anterior con respecto a la iteración actual.
A continuacion se mostrará los resultados de las pruebas del metodo usando la matriz Hessiana inicial de manera analítica y aproximada. Como tolerancia del gradiente se usó 10^{-3}.

Cuadro 1: Resultados de las pruebas realizadas al método BFGS con dos tipos de Hessiana

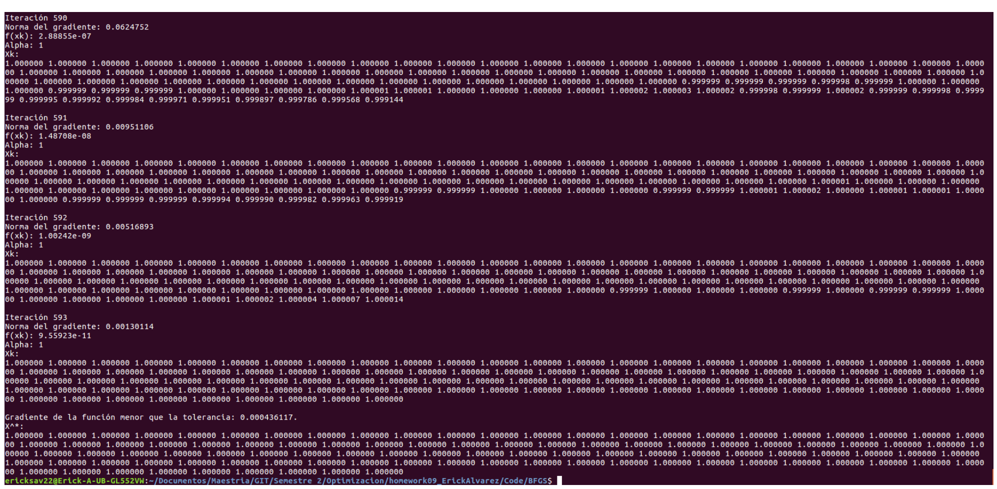
n	**h**	**Iteración final** k	$(x_k)_1$	$(x_k)_2$	$f(x_k)$	$\|\|\nabla f(x\_k)\|\|$
2	0.001	32	(0.999992, 0.999984)	(0.999999, 0.99998)	5.82602e-12	0.000102802
2	0.0001	34	(0.999992, 0.999984)	(0.999994, 0.999987)	1.30157e-10	0.000419161
2	0.00001	32	(0.999992, 0.999984)	(0.999976, 0.999954)	7.10816e-10	0.000598497
100	0.001	601	(1, 1, ..., 1)	(1, 1, ..., 0.999, 0.999)	3.79105e-11	0.000264983
100	0.0001	597	(1, 1, ..., 1)	(1, 1, 1.00003, 1.00005)	4.39948e-10	0.000861708
100	0.00001	593	(1, 1, ..., 1)	(1, 1, ..., 1)	9.55923e-11	0.000436117

En la Tabla anterior se puede apreciar los resultados de la ejecución del algoritmo, en donde se describe el número de la dimensión de la función, ahí solo se trabajó con las dimensiones 2 y 100, posteriormente se usaron tres valores de h para calcular la aproximación de la Hessiana que posteriormente se invertiría, luego se muestra el número de iteraciones que hizo el algoritmo para encontrar el óptimo. Aquí se añaden los que corresponden a la versión de la Hessiana aproximada ya que en general la otra versión tardó lo mismo o un poco menos que esta. Luego se muestran los vectores x_k resultantes del algoritmo, tanto de una versión como la otra, y finalmente se añade el valor de la función y la norma del gradiente en la última iteración.
En general se puede apreciar que el algoritmo trabajó muy bien con todas las pruebas que se hicieron, pudo encontrar los óptimos globales y el error en los mismos fue mínimo, esto hablando de las dos versiones.

A continuación se muestran dos capturas de pantalla las cuales fueron hechas en las últimas iteraciones del algoritmo con las dimensiones 2 y 100 y usando la versión aproximada de la Hessiana.



(a) Figura 1. Ejecución del algoritmo BFGS con dimensión 2.



(b) Figura 2. Ejecución del algoritmo BFGS con dimensión 100.

3. Notas y comentarios finales.

En el presente reporte se analizó la implementación del método BFGS y se mostraron los resultados de dos pruebas, unas con la matriz Hessiana inicial calculada de manera analítica y otra calculando la misma matriz usando el enfoque de diferencias finitas.
En ambos casos se apreció que el método trabajó muy bien, ya que este pudo encontrar el óptimo global de la función en relativamente pocas iteraciones, además vimos que el metodo no sé atasco en óptimos locales como otros metodos suelen hacerlo con la funcion de Rosenbrock.
Algo que hay que comentar sobre las pruebas con la aproximación de la Hessiana es que pese a que se usaron diferentes valores de h que en ocasiones generaban aproximaciones no tan buenas de la matriz, el metodo logró llegar al optimo de la función, aunque con más iteraciones. En esta parte logramos ver las propiedades autocorrectivas del metodo con puntos cercanos al optimo local.