Tarea 04 - Optimización

Giovanni Gamaliel López Padilla

Resumen

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Índice

1. Introducción

La optimización es procedimiento que con sus resultados se toman de decisiones en el análsis de sistemas físicos. Para realizar es la optimización de un sistema o situación se debe de contemplar un objetivo, el cual debe ser caracterizado por una función cuantitativa.

El resultado de realizar la optimización a un sistema puede representarse como un ahorro de tiempo, energía o cualquier objeto que pueda ser reflejado en un número. El objetivo del proceso de una optimización es obtener un conjunto de números o caracteristicas que representen un mínimo o máximo de objeto el cual esta siendo caracterizado. Esto puede ser representado como se encuentra en la ecuación 1.

$$\underset{x\in\mathbb{R}}{min}\ f(x) \tag{1}$$

donde f es la caracterización cuantitativa del problema y x son los objetos que interactuan con el sistema.
Una optimización local es una solución al problema de optimización en una vecindad alrededor del valor de x encontrado. En cambio una optimización global es aquella solución que es menor o mayor con respecto a todas las demás. La solución de un proceso de optimización no siempre encontrará los valores en que el sistema se situe una optimización global.

2. Marco teórico

2.1. Convexidad

El concepto de convexidad en los problemas de optimización es de gran importancia. Este concepto aporta una mayor facilidad al resolver un problema. La convexidad puede ser aplicado a conjunto o funciones. Se dice que S es un conjunto convexo si un segmento de linea conecta cualquier par de puntos en S. Sean $x, y \in S$, entonces $\alpha x + (1-\alpha)y \in S$ donde $\alpha \in [0, 1]$. Se dice que una función es convexa si su dominio S es convexo y cumplen con la propiedad escrita en la ecuación 2.

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \tag{2}$$

Donde $x, y \in S$ y $\alpha \in [0, 1]$.

2.2. Condiciones necesarias

Suponiendo que $f : \mathbb{R}^n \to \mathbb{R}$ y que es continuamente diferenciable y $p \in \mathbb{R}^n$. Entonces, tenemos que

$$f(x+p) = f(x) + \nabla f(x+tp)^T p$$

donde $t \in (0, 1)$. De igual manera, si f es doblemente continua diferenciable, entonces se tiene que:

$$\nabla f(x+p) = \nabla f(x) \int_0^1 \nabla^2 f(x+tp)pdt$$

por lo tanto

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla f(x+tp)p$$

Con estas hechos, se puede demostrar que si x^* es un punto estacionario entonces el gradiente y el hessiano de f tiene las caracteristicas mostradas en la ecuación 3.

$$\nabla f(x^*) = 0 \qquad \nabla^2 f(x^*) \succ 0 \tag{3}$$

También se puede demostrar que si $\nabla^2 f$ es continua en una vecindad alrededor de x^* y que $\nabla f(x^*) = 0$ y $\nabla^2 f(x^*)$ es positiva definida. Entonces x^* es un mínimo local de f.

2.3. Direcciones de búsqueda

La dirección mas eficiente usando un método de descenso es usar una dirección p descrita en la ecuación 4 en cada paso.

$$p_k = -\nabla f_k \tag{4}$$

En cada iteración de la linea de búsqueda se implementa un cambio en la posición siguiendo la ecuación 5.

$$x_{k+1} = x_k + \alpha_k p_k \tag{5}$$

Donde α_k es un escalar positivo llamado tamaño de paso.

2.4. Condiciones de Wolfe

Existen maneras de verificar si el α_k elegido para el k-esimo paso es el óptimo para seguir en la dirección p_k. La condición de decrecimiento suficiente esta descrita en la ecuación 6.

$$f(x_k + \alpha_k p_k) \leq f(x_k) c_1 \alpha_k \nabla f_k^T p_k \tag{6}$$

La interpretación de este resultado es que la función f debe ser proporcional al tamaño de paso α_k y la derivada direccional $\nabla f_k^T p_k$ para una constante $c_1 \in (0,1)$. Esta condición también es conocida como la condición de Armijo.

La condición de decrecimiento suficiente no es la única que se tiene que contemplar, esto debido a que existen α_k muy pequeñas que satisfacen a la desigualdad. Es por ello que se tiene que contemplar la condición de curvatura. La condición de curvatuva esta definida en la ecuación 7.

$$\nabla f(x_k + \alpha_k p_k)^T \geq c_2 \nabla f_k^T p_k \tag{7}$$

Donde $c_2 \in (0,1)$.

La búsqueda de un α_k que cumpla las condiciones de las ecuaciones 6 y 7 esta descrito en el algoritmo 1.

Algorithm 1: Búsqueda de un α que cumpla las condiciones de las ecuaciones 6 y 7

$\alpha_0 \leftarrow 0 \qquad \alpha_i \leftarrow 1 \qquad \beta \leftarrow \infty$
repeat
    if *armijo conditon(α_i) or curvature contidion(α_i)* **then**
        if *armijo condition(α_i)* **then**
            $\beta \leftarrow \alpha_i$
            $\alpha_i = \frac{\beta+\alpha}{2}$
            else if *curvature condition (α_i)* **then**
                $\alpha \leftarrow \alpha_i$
                if *β equals ∞* **then**
                    $\alpha_i \leftarrow 2\alpha$
                    else
                        $\alpha_i = \frac{\beta+\alpha}{2}$
    else
        break
return α_i

3. Métodos

3.1. Método de descenso del gradiente

La dirección que se tomada con el método del descenso del gradiente es el negativo del gradiente, esto es que la ecuación de cada paso en el método esta descrita en la ecuación 8.

$$x_{k+1} = x_k - \alpha_k \nabla f_k \tag{8}$$

donde se dice que α_k es fijo si para todo k el valor de α no cambia. Por lo que se convierte en un parametro del método.

3.2. Método de Newton

La direccion que es tomada en el método de Newton es obtenida al resolver el sistema matricial de la ecuación 9.

$$\nabla^2 f_k d_k = \nabla f_k \tag{9}$$

donde $-d_k$ es la dirección de descenso que se tomaria siguiendo la ecuación 10.

$$x_{k+1} = x_k - \alpha_k d_k \tag{10}$$

donde α_k sería obtenido en cada iteración siguiendo el algoritmo 1.

3.3. Condiciones de paro

3.3.1. Por posición

Una de las condiciones tomadas para parar los métodos de Newton y descenso del gradiente es comprobando si la norma de la diferencia entre la posición k y $k+1$ es menor igual a una tolerancia establecida. Esta condición es descrita en la ecuación 11.

$$||x_{k+1} - x_k|| \leq \tau \tag{11}$$

3.3.2. Por función

Otra manera de comprobar si ya no existe algún movimiendo notable en el método es revisando el valor absoluto de la diferencia de la función evaluada en x_{k+1} y x_k. La condición esta descrita en la ecuación 12.

$$|f(x_{k+1}) - f(x_k)| \leq \tau \tag{12}$$

3.3.3. Por gradiente

Como estamos en la búsqueda de puntos estacionarios de f, es razonable que una condición de paro este relacionada al gradiente de la function. La condición usando al gradiente de la función esta descrita en la ecuación 13.

$$||\nabla f(x_{k+1})|| \leq \tau \tag{13}$$

3.4. Algoritmos implementados

3.4.1. Método del descenso del gradiente

El algoritmo 2 se implemento para el método del descenso del gradiente. El algoritmo hace uso de las condiciones de paro descritas en 3.3.

Algorithm 2: Método del descenso de gradiente con α fija.

Input: x_0 , f , ∇f , α
Output: x_{i+1} $\nabla f(x_{i+1})$

$x_j \leftarrow x_0$
repeat
  $x_i \leftarrow x_j$
  $d_i \leftarrow -\nabla f(x_i)$
  $x_{i+1} \leftarrow x_i + \alpha d_i$
  check stop conditions$(f, x_i, x_{i+1}, \nabla f(x_{i+1}))$
return $x_{i+1}, \nabla f(x_{i+1})$

3.4.2. Método de Newton

El algoritmo 3 se implemento para el método de Newton. El mismo hace de las condiciones de Wolfe descritas en la sección 2.4 y las condiciones de paro descritas en 3.3.

Algorithm 3: Algoritmo de Newton usando las condiciones de Wolfe.

Input: x_0 , f , ∇f
Output: x_{i+1} $\nabla f(x_{i+1})$

$x_j \leftarrow x_0$
repeat
  $x_i \leftarrow x_j$
  $d_i \leftarrow$ solve$(\nabla^2 f(x_i) d_i = \nabla f(x_i))$
  $\alpha_i \leftarrow$ Wolfe conditions()
  $x_{i+1} \leftarrow x_i - \alpha_i d_i$
  check stop conditions$(f, x_i, x_{i+1}, \nabla f(x_{i+1}))$
return $x_{i+1}, \nabla f(x_{i+1})$

4. Resultados

5. Conclusiones

6. Referencias

[1] Ipiña A, López-Padilla G, Retama A, Piacentini RD, Madronich S. Ultraviolet Radiation Environment of a Tropical Megacity in Transition: Mexico City 2000–2019. Environmental Science & Technology. 2021 Aug;55(16):10946–10956. Available from: https://doi.org/10.1021/acs.est.0c08515.