Tarea 07 - Optimización

Giovanni Gamaliel López Padilla

Resumen

Los métodos del descenso del gradiente y de Newton son los métodos más simples en la rama de la optimización. Esto debido a que se basan en principios matemáticos los cuales indican que existe un mínimo global o local en una función con dominio convexo. En este trabajo se analizo la eficiencia de estos dos métodos usando las funciones de Rosembrock, de Wood y de suavizado. Llegando así a que el método del descenso del gradiente tiende a obtener mínimos locales cercanos al mínimo global en cuestión del valor de la función a analizar. Por otro lado el método de Newton obtiene un mínimo local cercano al punto inicial con un número menor de iteraciones a comparación del método del descenso del gradiente.

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

Una técnica para la optimización de funciones es definir la función de objetivo, generalmente para obtener alguna solución del problema se escogen pasos de búsqueda unidireccionales.

2.1. Región de confianza

Un método empleado para acotar las soluciones del problema es usar una región de confianza dado un punto. Para formalizar este método se define un modelo que aproxima a la función objetivo en un punto. El modelo se encuentra descrito en la ecuación 2.

$$m_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T \nabla^2 f(x_k) p \tag{2}$$

donde su mínimo se encuentra dentro de la región de confianza.
Une medida para calcular el radio de la región de confianza Δ_k es la medida de ajuste que esta definida en la ecuación 2.1.

$$\rho_k = \frac{f(x_k) - f(x + p_k)}{m_k(0) + m_k(p_k)}$$

Donde el numerador representa la reducción en la función y el denominador la reducción en el modelo. El denominador siempre es positivo debido a que p_k minimiza el modelo en cada iteración. En el caso en que ρ_k sea un valor negativo este representa un incremento en la función f. Por lo que el paso deberá rechazarse. Si ρ_k es muy cercano a uno, entonces el comportamiento de la función f y el modelo m_k tienen una semejanza, por lo que se optaría a incrementar la región de confianza. Si ρ_k se encuentra entre 0 y 1, entoces se opta por no realizar modificaciones en la región de confianza. Si ρ_k es cercano a cero o negativo se propone una reducción en la región de confianza.

2.2. Punto de Cauchy

Se denomina al arco de Cauchy al segmento tal que cumple la ecuación 3.

$$x_k^C(t) = \{x : x = x_k + t\nabla f(x_k), t \leq 0, ||t\nabla f(x_k)|| < \Delta_k\} \tag{3}$$

El punto de Cauchy se encuentra definido en la ecuación 4.

$$p_k^S = \underset{||p||<\Delta_k}{arg\ min}\ m_k(p) \tag{4}$$

donde Δ_k es la región de confianza.
Teniendo el punto p_k^S, se buscará un parámetro τ_k, el cual mínimize el modelo de la ecuación 2 en la región de confianza (ecuación 5).

$$\tau_k = \underset{\tau_k \geq 0}{arg\ min}\ m_k(\tau p_k^S) \leq \Delta_k \tag{5}$$

por lo que el punto de Cauchy, lo estaremos calculando de tal manera que $p_k^C = \tau_k p_k^S$.
Tomando a p_k como $-\lambda_k g_k$, sujeto a la región de confianza, se tiene que

$$||p|| = || - \lambda_k g_k|| \Rightarrow \lambda_k \leq \frac{\Delta_k}{||g_k||}$$

Por lo que la ecuación 5 se puede escribir como en la ecuación .

$$\lambda_k = \underset{\lambda \in [0,\hat{\lambda}]}{arg\ min}\ m(-\lambda g_l) \tag{6}$$

La solución al problema planteado en la ecuación 6 se encuentra descrita en la ecuación 7.

$$\lambda_k = \hat{\lambda} \begin{cases} 1 & \text{si } g_k^T B_k g_k \\ min\left(1, \frac{||g_k||^3}{\Delta_k g_k^T B_k g_k}\right) & \text{en cualquier otro caso} \end{cases} \tag{7}$$

Por lo tanto, el punto de Cauchy se obtiene como en la ecuación 8.

$$p_k^C = -\tau_k \frac{\Delta_k}{||g_k||} g_k \tag{8}$$

2.3. Método Dogleg

La idea del método Dogleg es minimizar el modelo cuadrático sin restricciones a lo largo del gradiente. La trayectoria Dogleg toma como primera instancia el punto p_k^U y como segunda linea va desde p_k^U hasta p_k^B. Lo anterior puede resumirse en la ecuación 9.

$$\check{p}(\tau) = \begin{cases} \tau p_k^U & \text{si } 0 \leq \tau \leq 1 \\ p_k^U + (\tau - 1)(p_k^B - p_k^U) & \text{si } 1 \leq \tau \leq 2 \end{cases} \tag{9}$$

Si se toma a p_k^U en dirección al gradiente, de tal manera que $p_k^U = \alpha \nabla f_k$, entonces el problema será encontrar α tal que p_k^U minimiza al modelo cuadrático (ecuación 2). Llegando a que α debe tener el valor descrito en la ecuación .

$$\alpha^* = -\frac{||\nabla f_k||}{\nabla^T f_k B_k \nabla f_k} \tag{10}$$

Por lo tanto, los puntos p_k^U y p_k^B son:

$$p_k^U = -\frac{||\nabla f_k||}{\nabla^T f_k B_k \nabla f_k} \nabla f_k \qquad p_k^B = -B_k^{-1} \nabla f_k$$

2.4. Método de Newton modificado

3. Métodos

3.1. Región de confianza

El método descrito en la sección 2.1 puede reducirse en el algoritmo 1.

Algorithm 1: Región de confianza

Input: $\hat{\Delta}$,$\Delta_0 \in (0, \hat{\Delta})$ y $\eta \in [0, \frac{1}{4}]$
for *k=0,1,2,3* **do**
  $p_k \leftarrow arg\ min$ ecuación 2
  $\rho_k \leftarrow$ ecuación 2.1
  if $\rho_k < \eta_1$ **then**
    $\Delta_{k+1} \leftarrow \hat{\eta}_1 \Delta_k$
    else if $\rho_k > \eta_2\ y\ ||p_k|| = \Delta_k$ **then**
      $\Delta_{k+1} \leftarrow min\{\hat{n}_2 \Delta_k, \hat{\Delta}\}$
      else
        $\Delta_{k+1} \leftarrow \Delta_k$

3.2. Método Dogleg con Región de confianza

El método descrito en la sección 2.3

Algorithm 2: Método de Dogleg con región de confianza

Input: $\hat{\Delta} > 0$, $\Delta_0 \in (0, \hat{\Delta})$ y $\eta \in [0, \frac{1}{4}]$
for $k = 0, 1, 2, 3, \ldots$ **do**
  $p_k \leftarrow$ ecuación 9
  $\rho_k \leftarrow$ ecuación 2.1
  if $\rho_k > \eta$ **then**
    $x_{k+1} \leftarrow x_k + p_k$ **else**
      $x_{k+1} = x_k$
  $\Delta_{k+1} \leftarrow$ algoritmo 1