Optimización I
Zaira Carolina Martinez Vargas

Tarea 8:
Métodos de región de confianza

2023-03-27

Resumen

1. Introducción

Los métodos de búsqueda lineal y los métodos de región de confianza generan tamaños de paso con un modelo lineal o cuadrático de la función objetivo, sin embargo, lo usan de forma distinta. En los métodos de región de confianza, primero definen una región y después escogen un tamaño de paso que minimice al modelo en esa región. Si el tamaño de paso no es adecuado, se reduce la región de confianza y se busca otro tamaño de paso. El tamaño de la región es importante porque si se escoge un tamaño pequeño, el algoritmo pierde la oportunidad de tomar un paso importante que se acerque más al mínimo de la función; en cambio, si es muy grande la región de confianza, el mínimo puede estar lejos y se tendría que reducir la región y volver a buscar. Así, se escoge el tamaño de la región de confianza dependiendo del rendimiento del algoritmo [1]. En este trabajo se utilizan algoritmos de región de confianza, para encontrar los mínimos de tres funciones: Rosenbrock (ec. 1), Wood (ec.2) y Branin (ec.3),

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

$$\begin{aligned} f(\mathbf{x}) = {} & 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 \\ & + 90(x_3^2 - x_4)^2 + 10{,}1\left[(x_2 - 1)^2 + (x_4 - 1)^2\right] + 19{,}8(x_2 - 1)(x_4 - 1), \end{aligned} \tag{2}$$

$$f(\mathbf{x}) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)cos(x_1) + s. \tag{3}$$

2. Métodos

2.1. Método Dogleg

Este método sólo se puede utilizar cuando el Hessiano de la función es positivo definido. En este se encuentra una solución aproximada reemplazando la curva de la trayectoria por otro término, con un camino que consiste de dos segmentos. La primera línea va del origen al minimizador con tamaño de paso del gradiente descendente, y la segunda línea va en función de una variable τ. De esta forma, se encuentra un p que minimiza el modelo a través de un camino, sujeto al borde de la región de confianza.

Algorithm 1 Algoritmo de Dogleg

Vector inicial x_0, gradiente de la función: g, Hessiano de la función: B,
for k = 0, 1, 2, ... **do**
  pU = -g^{T}Bg, pB = -B^{-1}g
  Resolver una ecuación cuadrática que depende de pU y pB: a = (pB-pU)^2; b = 2(pB^T)(pB-pU); c = $||pU||^2 - \Delta^2$
  tau1 = root1 + 1, tau2 = root2+1,
  if $0 \leq$ tau1 ≤ 2 **and** $0 \leq$ tau2 ≤ 2 **then**
    tau = min(tau1, tau2)
  else
    tau = tau2
  if tau ≤ 1 **then**
    pk = tau*pU
  else
    pk = pU+(tau-1)*(pB-pU)
  Reducir la región de confianza
  Calcular rho_k
  if $\rho > \eta$ **then**
    x = x + pk
  Recalcular Δ

2.2. Método de Newton-Cauchy

Este método alterna entre el paso de Newton y el paso de Cauchy, toma el paso de Newton si está en la región de confianza y en otro caso, toma el paso de Cauchy.

Algorithm 2 Newton-Cauchy

```
Vector inicial x_0, gradiente de la función: g, Hessiano de la función: B,
if Si B no es semidefinido positivo then
    tau = 1.0
else
    tau = min(1, ||g||^3/Δ * (g^T Bg))
pk = -1.0*(tau*Δ/||g|| )*g
Calcular ρ y utilizar las condiciones para región de confianza
```

2.3. Método de Newton Modificado

Este método, encuentra el paso como lo hace Newton, pero si el Hessiano de la función es singular, se modifica, en este caso, se añadió un término a los elementos de la diagonal para quitar la singularidad y se encontró el tamaño de paso usual: $pk = B^{-1}g$.

2.4. Retrospective filter trust-region (RFTR) [2]

Los pasos de este algoritmo son los siguientes:

1. Se inicializa como el método de la región de confianza básico, con un radio, y se inicializa un filtro F.
2. El modelo de la función es el mismo que el modelo de la región de confianza.
3. Para la primera iteración, se debe de calcular pk como se calcula en el algoritmo básico de región de confianza. En esta misma iteración, se calcula ρ y se cambia el punto:
   - Si $\rho > \eta_1$ entonces $x = x^+$
   - Si $\rho < \eta_1$ **and** $||g(x^+)|| < F$ entonces, $x = x^+$,
   - En otro caso, x no cambia.
4. A partir de la segunda iteración, se deben cumplir condiciones nuevas respecto a otra región de confianza.

3. Resultados

Cada algoritmo se ejecutó 30 veces y se calculó el error respecto al mínimo de la función que era conocido, para todos se inicializó con el punto mínimo más un valor aleatorio de una distribución uniforme en el rango de -2 a 2. En el Cuadro 1. se muestran los resultados con esta inicialización y en el Cuadro 2. con inicialización aleatoria. La línea que se muestra en los resultados significa que no funcionó el algoritmo y la norma del gradiente no se minimizaba.

	Dogleg	Newton-Cauchy Alterno	Newton Modificado	RFTR
Función Rosenbrock (n = 100)	0.3095	0.0996	0.0519	0.0302
Función Wood	0.4486	0.0039	0.0001	0.6344
Función Branin	0.7012	0.6184	0.7181	-

Cuadro 1: Inicialización: punto mínimo más un valor aleatorio.

	Dogleg	Newton-Cauchy Alterno	Newton Modificado	RFTR
Función Rosenbrock (n = 100)	0.0018	7.90e-7	0.1496	0.5214
Función Wood	0.1158	0.0016	0.0769	0.1389
Función Branin	0.6607	0.4332	1.8708	-

Cuadro 2: Inicialización aleatoria.

4. Conclusiones

Una ventaja de estos algoritmos, es que un mejor o peor resultado depende fuertemente del vector de inicialización, por ejemplo, en todos los algoritmos se utilizó el mismo valor de $\alpha = 1{,}0$, y en ninguno mostraba diferencias cambiando este parámetro. En los Cuadros 1. y 2. se observa que no hay ningún algoritmo que gane con las tres funciones, el algoritmo Newton modificado fue el mejor para la función Wood, pero el algoritmo Newton-Cauchy fue el mejor para la función Branin. Además, se observa que el algoritmo RFTR no funcionó con ninguna inicialización, y que la norma del gradiente no se disminuyó como se esperaba en las iteraciones.

Referencias

[1] Stephen Wright, Jorge Nocedal, et al. Numerical optimization. *Springer Science*, 35(67-68):7, 1999.

[2] Yue Lu, Zhongwen Chen, et al. A retrospective filter trust region algorithm for unconstrained optimization. *Applied Mathematics*, 1(03):179, 2010.