Orden del Tema

Punto de Cauchy

Definición

El Punto de Cauchy es el minimizador del modelo m_k a lo largo de la dirección del máximo descenso de la funcion f, i.e., $-\nabla f(x_k)$, sujeto a la región de confianza.

Alternativa para hallar el paso

La alternativa de solución del problema de optimización para hallar el paso recibe el nombre del Método Dogleg (Próxima clase) y está basada en el cáculo de el Punto de Cauchy .

Punto de Cauchy

- Para hallar el paso, se resuelve el problema de opimimización con restricciones:

$$\begin{aligned} p_k^* = \arg\min_p m_k(p) &= f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p, \\ \text{s.t.} & \quad \|p\| \leq \Delta_k. \end{aligned}$$

  Donde Δ_k es el radio de la región de confianza.
- Aunque en principio uno busca la solución del problema anterior, en la práctica, es suficiente encontrar una aproximación de p_k en la región de confianza que de un suficiente descenso del modelo para garantizar una convergencia del método.

Punto de Cauchy

- Para hallar el paso, se resuelve el problema de opimimización con restricciones:

$$
\begin{aligned}
p_k^* = \arg \min_p m_k(p) \quad &= \quad f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p, \\
\text{s.t.} \quad &\|p\| \leq \Delta_k.
\end{aligned}
$$

  Donde Δ_k es el radio de la región de confianza.

- Aunque en principio uno busca la solución del problema anterior, en la práctica, es suficiente encontrar una aproximación de p_k en la región de confianza que de un suficiente descenso del modelo para garantizar una convergencia del método.

Punto de Cauchy

- Para hallar el paso, se resuelve el problema de opimimización con restricciones:

$$
\begin{aligned}
p_k^* = \arg\min_p m_k(p) \quad &= \quad f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p, \\
\text{s.t.} \quad & \|p\| \leq \Delta_k.
\end{aligned}
$$

  Donde Δ_k es el radio de la región de confianza.

- El Punto de Cauchy, denotado como p_k^C, nos permite cuantificar el suficiente descenso del modelo.

Cálculo del Punto de Cauchy

Algoritmo: Punto de Cauchy

- Encontrar el punto p_k^S que resuelve la versión lineal:

$$p_k^S = \arg\min_p f(x_k) + \nabla f(x_k)^T p, \ \text{s.t.} \ \|p\| \leq \Delta_k$$

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \ \text{s.t.} \ \|\tau p_k^S\| \leq \Delta_k$$

- Calcular el Punto de Cauchy haciendo $p_k^C = \tau_k p_k^S$.

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 1)

- Encontrar el punto p_k^S que resuelve la versión lineal:

$$p_k^S = \arg\min_p f(x_k) + \nabla f(x_k)^T p, \ \text{s.t.} \ \|p\| \leq \Delta_k$$

- La función decrece a lo largo de $-\nabla f(x_k)^T$, luego $p_k^S = -\lambda \nabla f(x_k)$ con $\lambda > 0$
- Como $\|p_k^S\| \leq \Delta_k$, entonces $\lambda \leq \frac{\Delta_k}{\|\nabla f(x_k)\|}$
- El máximo descenso se obtiene para $\lambda = \frac{\Delta_k}{\|\nabla f(x_k)\|}$, por lo que $p_k^S = -\frac{\Delta_k}{\|\nabla f(x_k)\|} \nabla f(x_k)$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 1)

- Encontrar el punto p_k^S que resuelve la versión lineal:

$$p_k^S = \arg\min_p f(x_k) + \nabla f(x_k)^T p, \text{ s.t. } \|p\| \leq \Delta_k$$

- La función decrece a lo largo de $-\nabla f(x_k)^T$, luego $p_k^S = -\lambda \nabla f(x_k)$ con $\lambda > 0$
- Como $\|p_k^S\| \leq \Delta_k$, entonces $\lambda \leq \frac{\Delta_k}{\|\nabla f(x_k)\|}$
- El máximo descenso se obtiene para $\lambda = \frac{\Delta_k}{\|\nabla f(x_k)\|}$, por lo que $p_k^S = -\frac{\Delta_k}{\|\nabla f(x_k)\|} \nabla f(x_k)$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 1)

- Encontrar el punto p_k^S que resuelve la versión lineal:

$$p_k^S = \arg \min_p f(x_k) + \nabla f(x_k)^T p, \ \text{s.t.} \ \|p\| \leq \Delta_k$$

- La función decrece a lo largo de $-\nabla f(x_k)^T$, luego $p_k^S = -\lambda \nabla f(x_k)$ con $\lambda > 0$
- Como $\|p_k^S\| \leq \Delta_k$, entonces $\lambda \leq \frac{\Delta_k}{\|\nabla f(x_k)\|}$
- El máximo descenso se obtiene para $\lambda = \frac{\Delta_k}{\|\nabla f(x_k)\|}$, por lo que $p_k^S = -\frac{\Delta_k}{\|\nabla f(x_k)\|} \nabla f(x_k)$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 1)

- Encontrar el punto p_k^S que resuelve la versión lineal:

$$p_k^S = \arg\min_p f(x_k) + \nabla f(x_k)^T p, \text{ s.t. } \|p\| \leq \Delta_k$$

- La función decrece a lo largo de $-\nabla f(x_k)^T$, luego $p_k^S = -\lambda \nabla f(x_k)$ con $\lambda > 0$
- Como $\|p_k^S\| \leq \Delta_k$, entonces $\lambda \leq \frac{\Delta_k}{\|\nabla f(x_k)\|}$
- El máximo descenso se obtiene para $\lambda = \frac{\Delta_k}{\|\nabla f(x_k)\|}$, por lo que $p_k^S = -\frac{\Delta_k}{\|\nabla f(x_k)\|}\nabla f(x_k)$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
  - $\nabla f(x_k)^T B_k \nabla f(x_k) \leq 0$
  - $\nabla f(x_k)^T B_k \nabla f(x_k) > 0$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
  - $\nabla f(x_k)^T B_k \nabla f(x_k) \leq 0$
  - $\nabla f(x_k)^T B_k \nabla f(x_k) > 0$

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
  - Si $\nabla f(x_k)^T B_k \nabla f(x_k) \leq 0$ entonces $m_k(\tau_k p_k^S)$ decrece a lo largo de p_k^S , i.e., del $-\nabla f(x_k)$, y se toma a τ como el mayor valor posible, es decir $\tau = 1$.

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
    - Si $\nabla f(x_k)^T B_k \nabla f(x_k) \leq 0$ entonces $m_k(\tau_k p_k^S)$ decrece a lo largo de p_k^S , i.e., del $-\nabla f(x_k)$, y se toma a τ como el mayor valor posible, es decir $\tau = 1$.

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
  - Si $\nabla f(x_k)^T B_k \nabla f(x_k) > 0$ entonces $m_k(\tau_k p_k^S)$ es una cuadrática covexa en τ. Si el mínimo se alcanza en el interior de la región de confianza, entonces $\tau = \|\nabla f(x_k)\|^3 / (\Delta_k \nabla f(x_k)^T B_k \nabla f(x_k))$, en caso contrario la solución está en la frontera, $\tau = 1$ similar al caso anterior.

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Para hallar una fórmula cerrada para τ_k se consideran 2 casos:
    - Si $\nabla f(x_k)^T B_k \nabla f(x_k) > 0$ entonces $m_k(\tau_k p_k^S)$ es una cuadrática covexa en τ. Si el mínimo se alcanza en el interior de la región de confianza, entonces $\tau = \|\nabla f(x_k)\|^3 / (\Delta_k \nabla f(x_k)^T B_k \nabla f(x_k))$, en caso contrario la solución está en la frontera, $\tau = 1$ similar al caso anterior.

Cálculo del Punto de Cauchy

Algoritmo: Paso a Paso (Paso 2)

- Encontrar el parámetro $\tau_k > 0$ que minimiza $m_k(\tau_k p_k^S)$ en la región de confianza, i.e.,

$$\tau_k = \arg\min_{\tau \geq 0} m_k(\tau p_k^S), \text{ s.t. } \|\tau p_k^S\| \leq \Delta_k$$

- Resumiendo: $p_k^C = -\tau_k \frac{\Delta_k}{\|\nabla f_k\|} \nabla f_k$.

$$\tau_k = \begin{cases} 1, & \text{si } \nabla f_k^T B_k \nabla f_k \leq 0 \\ \min\left(1, \frac{\|\nabla f_k\|^3}{\Delta_k \nabla f_k^T B_k \nabla f_k}\right), & \text{e.o.c.} \end{cases}$$

Representación gráfica: Paso de Cauchy

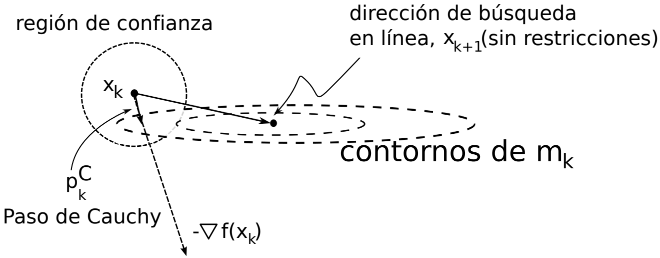

Figura: Punto de Cauchy

Punto de Cauchy: Otra forma de calcularlo

El Punto de Cauchy es el minimizador del modelo $m_k(p)$ a lo largo de de la dirección del máximo descendo, i.e., $p_k = -\lambda_k g_k$ sujeto a la región de confianza.

$$h(\lambda) \;\; := \;\; m_k(-\lambda g_k) = f_k - g_k^T g_k \lambda + \frac{1}{2}\lambda^2 g_k^T B_k g_k; \;\; \lambda \geq 0$$

Como $\|p\| \leq \Delta_k$ entonces

$$\|-\lambda g_k\| \leq \Delta_k \;\; \Rightarrow \lambda \leq \frac{\Delta_k}{\|g_k\|} =: \bar{\lambda}$$

$$\boxed{\lambda_k = \arg \min_{\lambda \in [0, \bar{\lambda}]} h(\lambda)}$$

Cálculo del Punto de Cauchy

La solucion del problema anterior es

$$
\begin{aligned}
\lambda_k &= \begin{cases} \bar{\lambda}, & \text{si } g_k^T B_k g_k \leq 0 \\ \min\left(\bar{\lambda}, \dfrac{\|g_k\|^2}{g_k^T B_k g_k}\right), & \text{e.o.c.} \end{cases} \\
&= \bar{\lambda} \begin{cases} 1, & \text{si } g_k^T B_k g_k \leq 0 \\ \min\left(1, \dfrac{\|g_k\|^3}{\Delta_k g_k^T B_k g_k}\right), & \text{e.o.c.} \end{cases} \\
&= \bar{\lambda}\tau_k
\end{aligned}
$$

Resumiendo: $p_k^C = -\lambda_k g_k$. Luego $p_k^C = -\bar{\lambda}\tau_k g_k = -\tau_k \frac{\Delta_k}{\|g_k\|} g_k$