Análisis Aplicado
Dr. Andreas Wachtel
Clase 14 (parte 1), 25 de Septiembre de 2019

Convergencia del método de región de confianza

Para obtener convergencia se requiere demostrar que la condición de aceptar un paso implica un descenso, es decir, tenemos que acotar

$$\rho_k \stackrel{\text{def}}{=\!=} \frac{f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})}{m(\mathbf{0}) - m(\boldsymbol{p}_k)} = \frac{\text{reducción real}}{\text{reducción del modelo}} \,.$$

Por lo cual, se requiere una cota inferior para el denominador. En general, se puede ver que cualquier punto $\boldsymbol{p}_k$ que es mejor o igual al punto de Cauchy implica una reducción suficiente para obtener convergencia.

Lema 1. *[Lemma 4.3, Nocedal] El punto de Cauchy satisface*

$$m(\mathbf{0}) - m(\boldsymbol{p}_C) \geq \frac{1}{2}\|\boldsymbol{g}_k\|_2 \cdot \min\left\{\Delta_k,\, \frac{\|\boldsymbol{g}_k\|_2}{\|B\|_2}\right\} \,.$$

Demostración. Hecho en clase (23 de Septiembre de 2019). □

Note que cualquier punto $\boldsymbol{p}_k$ que es mejor o igual que el punto de Cauchy, en el sentido $m(\boldsymbol{p}_k) \leq m(\boldsymbol{p}_C)$, satisface la misma condición, puesto que

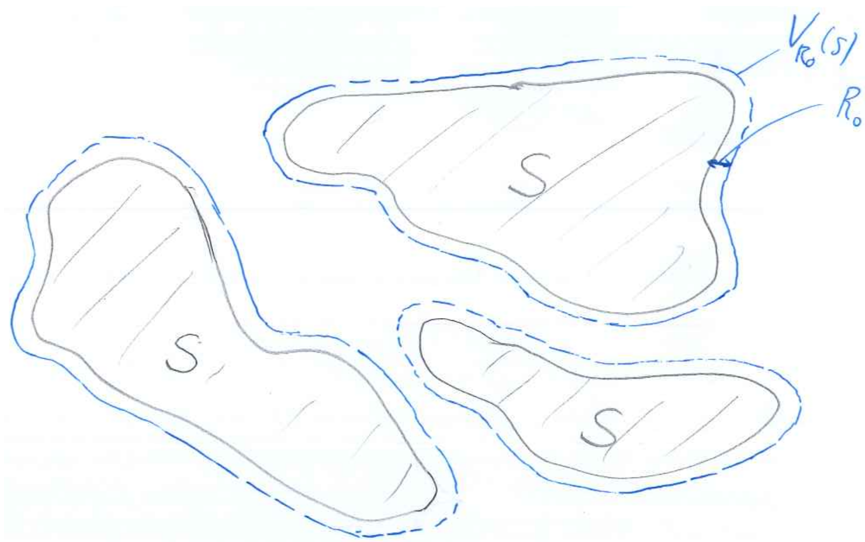
$$m(\mathbf{0}) - m(\boldsymbol{p}_k) \geq m(\mathbf{0}) - m(\boldsymbol{p}_C) \geq \frac{1}{2}\|\boldsymbol{g}_k\|_2 \cdot \min\left\{\Delta_k,\, \frac{\|\boldsymbol{g}_k\|_2}{\|B\|_2}\right\} \,. \tag{1}$$

Definición 1. Dado un punto inicial definimos el conjunto de *sub-nivel* $f(\boldsymbol{x}_0)$ por

$$S \stackrel{\text{def}}{=\!=} \{\boldsymbol{x} \in \mathbb{R}^n \colon f(\boldsymbol{x}) \le f(\boldsymbol{x}_0)\} \ ,$$

y una vecindad de S con distancia $R_0 > 0$ por

$$\mathcal{V}_{R_0}(S) \stackrel{\text{def}}{=\!=} \{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x} - \boldsymbol{y}\| < R_0 \quad \text{para algún} \quad \boldsymbol{y} \in S\} \ .$$



Teorema 2. *[Teorema 4.5, Nocedal] Suponga que* $\boldsymbol{x}^0 \in \mathbb{R}^n$ *y que*

1. $\eta = 0$, $\|B_k\| \le \beta$ *para todo* k,
2. f *es acotada por debajo en el conjunto de sub-nivel* S,
3. ∇f *es Lipschitz continua en* $\mathcal{V}_{R_0}(S)$,
4. *todo* $\boldsymbol{p}_k$ *satisface la cota* (1).

Entonces,

$$\liminf_{k\to\infty} \|\boldsymbol{g}_k\| = 0\,.$$

Demostración. Hicimos un bosquejo de la prueba de este teorema en clase 13 (23 de Septiembre). El bosquejo les debe ayudar a entender la prueba en el libro. □

Teorema 3. *[Teorema 4.6, Nocedal] Suponga que* $\boldsymbol{x}_0$, f, B_k *y* $\boldsymbol{p}_k$ *satisfacen las hipótesis del Teorema 2, pero con* $0 < \eta < \frac{1}{4}$. *Entonces,*

$$\lim_{k\to\infty} \|\boldsymbol{g}_k\| = 0\,.$$

Demostración. Sea $m > 0$ un índice tal que $\boldsymbol{g}_m \overset{\text{def}}{=\!=} \nabla f(\boldsymbol{x}_m) \neq \boldsymbol{0}$.
Como ∇f es Lipschitz continuo tenemos

$$\|\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{g}_m\| \leq L\|\boldsymbol{x} - \boldsymbol{x}_m\| \qquad \text{para todo} \qquad \boldsymbol{x} \in \mathcal{V}_{R_0}(S)\,.$$

Sean $\varepsilon \overset{\text{def}}{=\!=} \frac{1}{2}\|\boldsymbol{g}_m\|$ y $R \overset{\text{def}}{=\!=} \min\left\{\frac{\varepsilon}{L},\, R_0\right\}$.
Entonces, la bola $B_R(\boldsymbol{x}_m) \subset \mathcal{V}_{R_0}(S)$ y ∇f es Lipschitz en $B_R(\boldsymbol{x}_m)$.
Además, para todo $\boldsymbol{x} \in B_R(\boldsymbol{x}_m)$ tenemos

$$\|\boldsymbol{g}(\boldsymbol{x})\| \geq \|\boldsymbol{g}_m\| - \|\boldsymbol{g}_m - \boldsymbol{g}(\boldsymbol{x})\| \geq 2\varepsilon - L\|\boldsymbol{x} - \boldsymbol{x}_m\| \geq \varepsilon > 0\,.$$

Por lo cual, con Teorema 2 concluimos que la sucesión $\{\boldsymbol{x}_k\}_{k \geq m}$ tiene que salir de la bola $B_R(\boldsymbol{x}_m)$. Sea $\ell \geq m$ el primer índice tal que $\boldsymbol{x}_{\ell+1} \notin B_R(\boldsymbol{x}_m)$. Dado que $\|\boldsymbol{g}_k\| \geq \varepsilon$ para todo $k = m, m+1, \ldots, \ell$ sabemos que

$$\begin{aligned}
f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{\ell+1}) &= \sum_{k=m}^{\ell} f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1}) && \text{(suma telescópica)} \\
&\geq \sum_{k=m}^{\ell} \eta\Big(m_k(\boldsymbol{0}) - m_k(\boldsymbol{p}_k)\Big) && \text{(criterio de aceptar paso} \to \text{método)} \\
&\geq \sum_{k=m}^{\ell} \frac{\eta}{2}\,\|\boldsymbol{g}_k\|_2 \cdot \min\left\{\Delta_k,\, \frac{\|\boldsymbol{g}_k\|_2}{\|B\|_2}\right\} && \text{(por (1))} \\
&\geq \frac{1}{2}\eta\,\varepsilon \sum_{k=m}^{\ell} \min\left\{\Delta_k,\, \frac{\varepsilon}{\beta}\right\} && \text{(por } \|B_k\| \leq \beta \text{ y } \|\boldsymbol{g}_k\| \geq \varepsilon)\,.
\end{aligned}$$

Falta analizar la suma. Si $\Delta_k \leq \frac{\varepsilon}{\beta}$ para todo $k = m, m+1, \ldots, \ell,$ entonces $\sum_{k=m}^{\ell} \Delta_k \geq R,$ pues $\boldsymbol{x}_{\ell+1}$ salio de la bola $B_R(\boldsymbol{x}_m)$. En este caso, tenemos

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{\ell+1}) \geq \frac{1}{2}\eta\,\varepsilon\, R = \frac{1}{2}\eta\,\varepsilon\, \min\left\{\frac{\varepsilon}{L},\, R_0\right\}\,.$$

Por otro lado, si existe $\Delta_k > \frac{\varepsilon}{\beta},$ entonces

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{\ell+1}) \geq \frac{1}{2}\eta\,\varepsilon\,\frac{\varepsilon}{\beta}\,.$$

Unimos esos dos casos para concluir que

$$f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{\ell+1}) \geq \frac{1}{2}\eta\,\varepsilon\, \min\left\{\frac{\varepsilon}{\beta},\, \frac{\varepsilon}{L},\, R_0\right\}\,.$$

Finalmente, la sucesión $f(\boldsymbol{x}_k)$ es decreciente y acotada por debajo por $f_\star$. En particular, sabemos que $f(\boldsymbol{x}_m) \downarrow f(\boldsymbol{x}_\star)$ y con $2\varepsilon = \|\boldsymbol{g}_m\|$ concluimos que

$$f(\boldsymbol{x}_m) - f_\star \geq f(\boldsymbol{x}_m) - f(\boldsymbol{x}_{\ell+1}) \geq \frac{1}{4}\eta\,\|\boldsymbol{g}_m\|\, \min\left\{\frac{\|\boldsymbol{g}_m\|}{2\beta},\, \frac{\|\boldsymbol{g}_m\|}{2L},\, R_0\right\} \geq 0\,.$$

Es decir, $f(\boldsymbol{x}_m) - f_\star \downarrow 0 \implies \|\boldsymbol{g}_m\| \downarrow 0.$ $\square$