Local Error Bound (cont'd)

2022年3月7日 17:38

Recall the Setting of Scenario 2:

$$\hat{\theta} \in \underset{\theta \in \mathbb{R}^d}{\operatorname{argmin}} \left\{ F(\theta) \triangleq \underbrace{h}_{\substack{\text{strongly convex} \\ \text{on compact sets,} \\ \nabla h \text{ Lipschitz}}}(\overset{\mathbb{R}^{n\times d}}{A}\theta) + \underbrace{R(\theta)}_{\substack{\text{norm with} \\ \text{polyhedral} \\ \text{epigraph}}} \right\} \quad — (*)$$

As before, $\hat{v}$ is the optimal value and Θ is the optimal solution set.

Last lecture, we showed that

(1) (decomposition of Θ) there exists a $\bar{y} \in \mathbb{R}^n$ s.t. with $\bar{g} = A^T \nabla h(\bar{y})$, $\overset{A\hat{\theta},\ \hat{\theta}\in\Theta}{\shortparallel}$

$$\Theta = \underbrace{\{\theta \in \mathbb{R}^d : A\theta = \bar{y}\}}_{\Theta_L :\ \text{polyhedral}} \cap \underbrace{\{\theta \in \mathbb{R}^d : -\bar{g} \in \partial R(\theta)\}}_{\Theta_R = (\partial R)^{-1}(-\bar{g}) :\ \text{Polyhedral}},$$

(2) (linear regularity) there exists a constant $c > 0$ s.t.

$$\text{dist}(\theta, \Theta) \le c \cdot [\text{dist}(\theta, \Theta_L) + \text{dist}(\theta, \Theta_R)], \quad \forall \theta \in \mathbb{R}^d,$$

(3) (Hoffman error bound) there exists a constant $c' > 0$ s.t.

$$\text{dist}(\theta, \Theta_L) \le c' \cdot \|A\theta - \bar{y}\|_2, \quad \forall \theta \in \mathbb{R}^d.$$

Next tasks:

① The RHS of the error bound is in terms of $\|E(\theta)\|_2$. Hence, we need to relate $\|A\theta - \bar{y}\|_2$ to $\|E(\theta)\|_2$.

② We cannot apply the same trick to bound $\text{dist}(\theta, \Theta_R)$, since we do not have an explicit description of Θ_R.

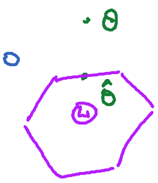
Thus, some new tools are needed.

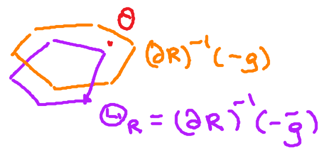
Let us tackle Task 2 first. Recall that

$$\Theta_R = \{\theta \in \mathbb{R}^d : -\bar{g} \in \partial R(\theta)\}$$

and for any $\hat{\theta} \in \Theta$, we have $\hat{\theta} \in \Theta_R$. Say $\theta \notin \Theta_R$

<u>Idea</u>: Suppose that $\mathrm{dist}(\theta, \Theta)$ is small. Let $\hat{\theta} = \Pi_{\Theta}(\theta)$, so that $\mathrm{dist}(\theta, \Theta) = \|\theta - \hat{\theta}\|_2$. Let $g = A^T \nabla h(A\theta)$, recall $\bar{g} = A^T \nabla h(\underbrace{A\hat{\theta}}_{=\bar{y}})$. Then, $\|g - \bar{g}\|_2 \leq L \cdot \|A\|^2 \cdot \|\theta - \hat{\theta}\|_2$ by Lipschitz continuity of ∇h and should be small. Now, Consider the set-valued map $u \mapsto (\partial R)^{-1}(-u)$. Note that $\Theta_R = (\partial R)^{-1}(-\bar{g})$. Since g is close to $\bar{g}$, we expect that $(\partial R)^{-1}(-\bar{g})$ is close to $(\partial R)^{-1}(-g)$, with the closedness measured in terms of $\|g - \bar{g}\|_2$.





If in addition $\theta \in (\partial R)^{-1}(-g)$, then we can bound $\mathrm{dist}(\theta, \Theta_R) = \mathrm{dist}(\theta, (\partial R)^{-1}(-\bar{g}))$ in terms of $\|g - \bar{g}\|_2$.

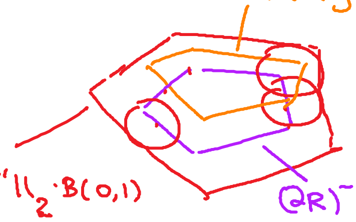
This motivates the need for a notion of Lipschitz continuity for set-valued maps, Something like

$$\text{"}\|(\partial R)^{-1}(-g) - (\partial R)^{-1}(-\bar{g})\| \leq \beta \cdot \|g - \bar{g}\|_2\text{"}$$

<u>Proposition 2</u>: (Outer Lipschitz Continuity of $(\partial R)^{-1}$)

There exists a $\beta > 0$ s.t. for any $g' \in \mathbb{R}^d$, there is a neighborhood $V_{g'}$ of g' s.t

$$\forall g'' \in V_{g'} : \quad (\partial R)^{-1}(g'') \subseteq (\partial R)^{-1}(g') + \beta \|g' - g''\|_2 \cdot B(0,1)$$

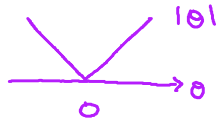


Recall:
$A + B = \{x + y : x \in A, y \in B\}$

In particular, any point in $(\partial R)^{-1}(g'')$ is at most at a distance of $\beta \cdot \|g' - g''\|_2$ from $(\partial R)^{-1}(g')$.

Example: Consider $R(\theta) = |\theta|$. Then, we have

$$\partial R(\theta) = \begin{cases} 1 & \text{if } \theta > 0 \\ [-1,1] & \text{if } \theta = 0 \\ -1 & \text{if } \theta < 0 \end{cases}$$

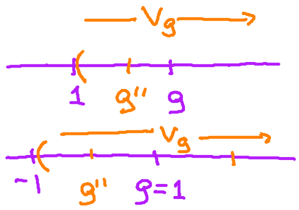


It follows that

$$(\partial R)^{-1}(g) = \begin{cases} \phi & \text{if } |g| > 1 \\ \mathbb{R}_+ & \text{if } g = 1 \\ \mathbb{R}_- & \text{if } g = -1 \\ \{0\} & \text{if } |g| < 1 \end{cases}$$

$$\shortparallel$$

$$\{\theta : g \in \partial R(\theta)\}$$



$\underset{\in(-1,1)}{(\partial R)^{-1}(g'')} = \{0\} \subseteq \mathbb{R}_+$

$(\partial R)^{-1}(1) = \mathbb{R}_+ \subseteq \mathbb{R}_+$

$(\partial R)^{-1}(\underset{>1}{g''}) = \phi \subseteq \mathbb{R}_+$

$(\partial R)^{-1}(-1) = \mathbb{R}_- \nsubseteq \mathbb{R}_+$

In particular, $(\partial R)^{-1}$ is OLC w/ $\beta = 0$

Using Proposition 2, we have

$$\operatorname{dist}(\theta, \underbrace{\Theta_R}_{\substack{\shortparallel \\ (\partial R)^{-1}(-\bar{g}) \\ \text{OLC}}}) \leq \beta \|g - \bar{g}\|_2 \quad \text{for any } \theta \in (\partial R)^{-1}(-g),\ -g \in V_{-\bar{g}}.$$

Hence, we have

$$\operatorname{dist}(\theta, \Theta) \leq \beta' \left[\|A\theta - \bar{y}\|_2 + \|g - \bar{g}\|_2 \right] \quad \text{— (B)}$$

for any $\theta \in (\partial R)^{-1}(-g)$, $-g \in V_{-\bar{g}}$.

Idea: $\|g - \bar{g}\|_2 \text{ "=" } \|A^T \nabla h(A\theta) - A^T \nabla h(\bar{y})\|_2$

$\leq \|A\| \cdot \|\nabla h(A\theta) - \nabla h(\bar{y})\|_2$

$\leq L \cdot \|A\| \, \|A\theta - \bar{y}\|_2$ by Lipschitz Continuity of ∇h.

Caution: $\theta \in (\partial R)^{-1}(-g)$, but it is not clear $g = A^T \nabla h(A\theta)$.

Roadmap:

① Observe:

$\|E(\theta)\|_2$

① Observe:

$$\mathrm{dist}(\theta, \Theta) \leq \mathrm{dist}(\theta + E(\theta), \Theta) + \|E(\theta)\|_2$$

② Use (B) to bound $\mathrm{dist}(\theta + E(\theta), \Theta)$

in terms of $\|E(\theta)\|_2$