4月10日

2020年4月10日 7:46

$$g \in \partial f(x) \iff f(y) \geq f(x) + g^T(y-x),\ \forall y \in \mathrm{dom} f.$$

Eg: $f(x) = \|x\|$

$$\partial f(x) = \{ y \mid \|y\|_* \leq 1,\ \langle y, x\rangle = \|x\| \}$$

$\Downarrow$

① If $f(x) = \|x\|_2 \Rightarrow \partial\|x\|_2 = \{ y \mid \|y\|_2 \leq 1,\ \langle y, x\rangle = \|x\|_2 \}$

$$= \begin{cases} \frac{x}{\|x\|_2}, & \text{if } x \neq 0 \\ \{y \mid \|y\|_2 \leq 1\}, & \text{if } x = 0. \end{cases}$$

Proof: $\forall g \in \partial f(x)$

$$\underline{\|y\| \geq \|x\| + \langle g, y-x\rangle} \quad (\triangle)$$

(1) $y = 2x \Rightarrow \|x\| \geq \langle g, x\rangle$

(2) $y = 0 \Rightarrow 0 \geq \|x\| + \langle g, -x\rangle \Rightarrow \langle g, x\rangle \geq \|x\|$

$\Rightarrow \langle g, x\rangle = \|x\|.$

$(\triangle) \iff \|y\| \geq \langle g, y\rangle \iff \|g\|_* \leq 1$

$X \in S_+^n$, $\quad \|X\|_* = \sum_i \sigma_i = \sqrt{Tr(X^T X)}$

Dual of nuclear norm $\iff$ Operator norm

$$\partial\|X\|_* = \{ Y \mid \|Y\|_2 \leq 1,\ \langle Y, X\rangle = \|X\|_* \}$$

$h(x,y) = \|x-y\|$, $f(x) = dist(x, C) = \inf_{y \in C} \|x-y\|_2$

$g \in \partial f(\hat{x})$ if $(g, 0) \in \partial h(\hat{x}, \hat{y})$, $\hat{y} = \arg\min_{y \in C} \|\hat{x} - y\|_2$

$$\partial_x h(\hat{x}, \hat{y}) = \partial_x \underbrace{\|\hat{x} - \hat{y}\|_2}_{>0} = \frac{1}{\|\hat{x} - \hat{y}\|_2}(\hat{x} - \hat{y}) \quad \left(\partial\|x\|_2 = \frac{x}{\|x\|_2} \text{ if } x \neq 0\right)$$

f: Strongly convex (μ)

GD: $O(\frac{1}{k}) \to$ Linear convergence

SD: $O(\frac{1}{\sqrt{k}}) \to O(\frac{1}{k})$

Proof: $$\|x^{k+1}-x^*\|_2^2 = \|x^k - x^*\|_2^2 - \underline{2t_k g^{kT}(x^k - x^*)} + t_k^2\|g^k\|_2^2$$

$$f^* \geq f^k + g^k(x^k - x^*) + \frac{\mu}{2}\|x^k - x^*\|_2^2$$

$$\leq (1-\mu t_k)\|x^k - x^*\|_2^2 - 2t_k(f^k - f^*) + t_k^2\|g^k\|_2^2$$

Choose $t_k = \dfrac{2}{\mu(k+1)}$

$$\Rightarrow k(f^k - f^*) \leq \frac{1-\mu t_k}{2t_k}\|x^k - x^*\|_2^2 - \frac{1}{2t_k}\|x^{k+1}-x^*\|^2 + \frac{t_k}{2}\|g^k\|^2$$

$$\leq \underbrace{\frac{\mu k(k-1)}{4}\|x^k - x^*\|_2^2}_{\delta} - \underbrace{\frac{\mu k(k+1)}{4}\|x^{k+1}-x^*\|^2}_{\Delta} + \frac{k\|g^k\|^2}{\mu(k+1)}$$

$$\leq \frac{\|g^k\|^2}{\mu}$$

$$\Rightarrow \underline{\sum_{j=0}^{k} j(f^j - f^*)} \leq \underbrace{-\frac{\mu k(k+1)}{4}\|x^{k+1}-x^*\|^2}_{\leq 0} + \sum_{j=0}^{k}\frac{1}{\mu}\underbrace{\|g^k\|^2}_{\leq G}$$

$$\sum_{j=0}^{k} j\,(f_{best}^{(k)} - f^*)$$

$O(k^2)$

$$\leq \frac{k}{\mu}G^2$$

$O(k)$

$$\Rightarrow f_{best}^{(k)} - f^* \leq \frac{2G^2}{\mu(k+1)}$$