

Theorem 1

For all $x \in Q$

1. Unbiased:

$$\mathbb{E}_e\{\nabla f_\gamma(x, e)\} = \nabla f(x), \gamma \rightarrow 0;$$

2.

$$\forall \gamma, \mathbb{E}_e\{\nabla f_\gamma(x, e)\} = \nabla f(x) + \mathbb{E}_e\{e\}(O(\gamma) + \frac{d\Delta}{\gamma})$$

3. Bounded variance:

$$\begin{aligned} \mathbb{E}_e\{\|\nabla f_\gamma(x, e)\|_q^2\} &= \kappa(p, d)(dM_2^2 + \frac{d^2\Delta^2}{\gamma^2}), \\ 1/p + 1/q &= 1, \\ \kappa(p, d) &= O(\sqrt{\mathbb{E}_e\|e\|_q^4}) = \begin{cases} O(1), p = 2; \\ O((\ln d)/d), p = 1. \end{cases} \end{aligned}$$

Proof

First two statements

Due to the existence of first derivative:

$$\begin{aligned} \vec{g} &:= \mathbb{E}_e\{\nabla f_\gamma(x, e)\} = \\ d\mathbb{E}_e\left\{\frac{f(x) + \langle \nabla f(x), \gamma e \rangle + o(\|\gamma e\|_2) - (f(x) - \langle \nabla f(x), \gamma e \rangle + o(\|\gamma e\|_2))}{2\gamma}\right\} e &= \\ = d\mathbb{E}_e\left\{\frac{2\langle \nabla f(x), \gamma e \rangle + o(|\gamma|)}{2\gamma}\right\} e &\underset{\gamma \rightarrow 0}{=} d\mathbb{E}_e\{\langle \nabla f(x), e \rangle e\} = \nabla f(x) \end{aligned}$$

Let $\delta(x) : |\delta(x)| \leq \Delta$ Oracle's noise, then:

$$\begin{aligned} \vec{g} &:= \mathbb{E}_e\{\nabla f_\gamma(x, e)\} = \\ d\mathbb{E}_e\left\{\frac{\delta(x + \gamma e) + \langle \nabla f(x), \gamma e \rangle - (\delta(x - \gamma e) - \langle \nabla f(x), \gamma e \rangle) + O(\|\gamma e\|_2^2)}{2\gamma}\right\} e &\leq \\ \leq d\mathbb{E}_e\left\{\frac{2\langle \nabla f(x), \gamma e \rangle + 2\Delta + O(\gamma^2)}{2\gamma}\right\} e &= \nabla f(x) + \mathbb{E}_e\{e\}O(d\gamma + \frac{d\Delta}{\gamma}), \end{aligned}$$

More than that, let $\vec{r} = x_0 - x^*, R = \|\vec{r}\|_2$

$$\begin{aligned} |\langle \vec{g}, \vec{r} \rangle| &= |\langle \nabla f(x), \vec{r} \rangle + \langle \mathbb{E}_e\{e\}, \vec{r} \rangle(dO(\gamma) + \frac{d\Delta}{\gamma})| \leq |\langle \nabla f(x), \vec{r} \rangle + \mathbb{E}_e\{\langle e, \vec{r} \rangle\}O(d\gamma + \frac{d\Delta}{\gamma})| \leq \\ &\leq |\langle \nabla f(x), \vec{r} \rangle| + \frac{R}{\sqrt{d}}|O(\gamma + \frac{\Delta}{\gamma})| \end{aligned}$$

Then maximal residual (for $\gamma = \sqrt{d\Delta}$):

$$\varepsilon \approx \frac{R}{\sqrt{d}} \cdot \sqrt{d\Delta} = R\sqrt{\Delta}$$

The third statement

Due to [Shamir, 2017](#) (Lemma 4 and 5) we obtain:

$$\begin{aligned}\mathbb{E}_e\{\|\nabla f_\gamma(x, e)\|_q^2\} &= \kappa(p, d)(dM_2^2 + \frac{d^2\Delta^2}{\gamma^2}), \\ 1/p + 1/q &= 1, \\ \kappa(p, d) &= O(\sqrt{\mathbb{E}_e\|e\|_q^4}) = \begin{cases} O(1), p = 2; \\ O((\ln d)/d), p = 1. \end{cases}\end{aligned}$$

Conclusion

For reaching ε sub-optimality the noise should be not greater then:

a)

$$R\sqrt{\Delta} \leq \varepsilon \implies \Delta \leq \frac{\varepsilon^2}{R^2}$$

b)

$$\begin{aligned}\frac{d^2\Delta^2}{\gamma^2} \leq dM^2 &\implies \Delta \leq \frac{\gamma M}{\sqrt{d}} = \lceil \gamma = \sqrt{d\Delta} \rceil = M \\ \Delta &\leq M^2\end{aligned}$$

So $\Delta \leq \min\{M^2, \frac{\varepsilon^2}{R^2}\} \approx \frac{\varepsilon^2}{R^2}$