Оценки максимально допустимого враждебного шума для задач безградиентной гладкой выпуклой оптимизации

Павлов Игорь Николаевич

Московский физико-технический институт
Физтех-школа прикладной математики и информатики

18 марта 2025 г.

Научный руководитель: д-р физ.-мат. наук, проф. Гасников А.В.

Контактная информация: pavlov.in@phystech.edu

Содержание

Постановка задачи

Рассмотрим задачу выпуклой оптимизации:

$$\min_{x \in Q} f(x)$$

Предполагая, что доступен только оракул 0-ого порядка (возвращающий значение функции) с некоторым абсолютным шумом.

$$O_f(x, \xi) = f(x) + \xi, \text{ где } |\xi| \leq \Delta$$

Необходимо найти точку x, такую что $f(x) - f(x_*) \leq \varepsilon$, где x_* - решение задачи.

Случай липщицевой целевой функции

В рамках работы [1] было доказано:

Theorem 1

For every constant $c \geq 1$, there exists a constant d_c such that for every algorithm A, every $d \geq d_c$, there exists a convex set $K \subseteq \mathbb{R}^d$ with diameter 1, a Δ-approximate convex function $\hat{f} : K \to \mathbb{R}$, and $\epsilon \in [0, 1/64]$ such that

$$\Delta \geq \max\left\{\frac{\epsilon^2}{\sqrt{d}}, \frac{\epsilon}{d}\right\} \times \left(13c \log \frac{d}{\epsilon}\right)^2,$$

and such that A fails to output, with probability $\geq 1/2$, a point $\tilde{x} \in K$ with

$$\hat{f}(\tilde{x}) \leq \min_{x \in K}\{\hat{f}(x)\} + \epsilon$$

in $o\left(\left(\frac{d}{\epsilon}\right)^c\right)$ time.

Случай липщицевой целевой функции

А также:

Theorem 2

Let d be a positive integer, $\delta > 0$ be a positive real number, ϵ, Δ be two positive real numbers such that

$$\Delta \le \max\left\{\frac{\epsilon^2}{\mu\sqrt{d}}, \frac{\epsilon}{d}\right\} \times \frac{1}{16348}.$$

Then there exists an algorithm $\mathcal{A}$ such that, given any Δ-approximate convex function $\tilde{f}$ over a μ-rounded convex set $\mathcal{K} \subseteq \mathbb{R}^d$ of diameter 1, $\mathcal{A}$ returns a point $\tilde{x} \in \mathcal{K}$ with probability $1 - \delta$ in time *poly* $\left(d, \frac{1}{\epsilon}, \log\frac{1}{\delta}\right)$ such that

$$\tilde{f}(\tilde{x}) \le \min_{x\in\mathcal{K}} \tilde{f}(x) + \epsilon.$$

Случай липщицевой целевой функции

Что приводит нас к следующей верхней оценке на максимально допустимый уровень шума:

$$\Delta = \tilde{O}\left(\max\left\{\frac{\varepsilon^2}{\sqrt{d}MR}, \frac{\varepsilon}{d}\right\}\right)$$

Случай липшицевой, сильно-выпуклой целевой функции

Теперь к результатам, полученным в статье [2].
Используя регуляризацию (строим функцию $f'(x) = f(x) + \frac{\mu}{2}\|x - x_*\|_2^2$) для сведения выпуклого случая к сильно выпуклому получаем оценку на максимально допустимый уровень шума:

$$\Delta = \tilde{O}\left(\max\left\{\frac{\mu^{1/2}\varepsilon^{3/2}}{\sqrt{d}M}, \frac{\varepsilon}{d}\right\}\right)$$

Случай гладкой целевой функции

Теперь используем сглаживание (строим функцию $f_\gamma(x) = \mathbb{E}_e[f(x + \gamma e)]$, где $u \sim \mathcal{U}(B_2^d(1))$, которая будет $\frac{\sqrt{d}M}{\gamma}$-гладкой) и получаем верхнюю оценку на максимально допустимый уровень шума в данном классе:

$$\Delta = \tilde{O}\left(\max\left\{\frac{\varepsilon^{3/2}}{\sqrt[4]{d}\sqrt{LR}}, \frac{\varepsilon}{d}\right\}\right)$$

Случай гладкой, сильно-выпуклой целевой функции

Комбинируя два использованных ранее подхода получаем верхнюю оценку на максимальный допустимый уровень шума в данном случае:

$$\Delta = \tilde{O}\left(\max\left\{\frac{\varepsilon^{3/2}}{\sqrt[4]{d}\sqrt{LR}}, \frac{\varepsilon}{d}\right\}\right)$$

Возможные дальнейшие направления развития

Известная нижняя оценка на максимальный допустимый уровень шума для гладкого случая ([3]):

$$\Delta \lesssim \frac{\varepsilon^{3/2}}{R\sqrt{dL}}$$

т. е. меньше, чем то, что получено в [2]. Возникает вопрос, можно ли улучшить оценку?

Литература

- Risteski, Andrej and Li, Yuanzhi. Algorithms and matching lower bounds for approximately-convex optimization. Advances in Neural Information Processing Systems, 2016. https://proceedings.neurips.cc/paper_files/paper/2016/file/186fb23a33995d91ce3c2212189178c8-Paper.pdf
- Dmitrii A. Pasechnyuk and Aleksandr Lobanov and Alexander Gasnikov. Upper bounds on the maximum admissible level of noise in zeroth-order optimisation. 2023. https://arxiv.org/abs/2306.16371
- Gasnikov, Alexander and Dvinskikh, Darina and Dvurechensky, Pavel and Gorbunov, Eduard and Beznosikov, Aleksandr and Lobanov, Alexander. Randomized Gradient-Free Methods in Convex Optimization. Encyclopedia of Optimization, 2023. https://arxiv.org/abs/2211.13566