Mixed Integer Programming

Mixed Integer Programming

Zero order methods. Gradient free optimization. Global optimization

- Шпаргалка по результатам в безградиентной оптимизации
- RL и эволюционные алгоритмы
- Open in Colab Global optimization illustration
- Open in Colab Nevergrad library
- Open in Colab Optuna quickstart
- Демонстрация медленности методов нулевого порядка
- Open in Colab Подбор гиперпараметров модели машинного обучения в Keras с помощью Optuna
- A Tutorial on Zero-Order Optimization

Case 1: 2-Point & Multi-Point Estimators

- A naïve approach:

$$\mathsf{G}_f^{2n}(x;u) = \sum_{i=1}^{n} \frac{f(x+ue_i) - f(x-ue_i)}{2u} e_i$$

- When f is L-smooth, we have

$$\|\mathsf{G}_f^{2n}(x;u) - \nabla f(x)\| \le \frac{1}{2} uL\sqrt{n}$$

where $f^u(x) = \mathbb{E}_{y\sim\tilde{\lambda}}[f(x+uy)]$ is a smooth version of f

$\lambda = \mathrm{Uni}(\mathbb{B}_n)$

Case 1: 2-Point & Multi-Point Estimators

- 2-point gradient estimator:

$$\mathsf{G}_f^{(2)}(x;u,z) = n\,\frac{f(x+uz) - f(x-uz)}{2u} z \qquad z \sim \lambda$$

where λ is spherically symmetric with $\mathbb{E}_{z\sim\lambda}\left[\|z\|^2\right] = 1$

$f^u = \mathbb{E}_z G_f^{(2)}$

- Some facts for L-smooth / convex / μ-strongly convex function f:
  - f^u is L-smooth / convex / μ-strongly convex
  - $|f^u(x) - f(x)| \le \frac{1}{2}u^2 L \cdot \frac{n}{n+2}\mathbb{E}_{z\sim\lambda}\left[\|z\|^4\right]$

    $\|\nabla f^u(x) - \nabla f(x)\| \le uL \cdot \frac{n}{n+1}\mathbb{E}_{z\sim\lambda}\left[\|z\|^3\right]$

$$\mathbb{E}_{z\sim\lambda}\left[\left\|\mathsf{G}_f^{(2)}(x;u,z)\right\|^2\right] \le (1+\kappa)\mathbb{E}_{z\sim\lambda}\left[\|z\|^4\right]\cdot n\|\nabla f(x)\|^2 + \frac{1+\kappa}{4\kappa}\mathbb{E}_{z\sim\lambda}\left[\|z\|^6\right]\cdot n^2u^2L^2$$