Online Newton Method for Bandit Convex Optimisation

Hidde Fokkema — University of Amsterdam, Korteweg de Vries Institute for Mathematics

Dirk van der Hoeven — Universiteit Leiden, Mathematical Institute

Tor Lattimore — Google DeepMind

Jack J. Mayo — University of Amsterdam, Korteweg de Vries Institute for Mathematics

Summary

We introduce a computationally efficient algorithm for zeroth-order bandit convex optimisation which achieves a **regret** of **at most** $d^{3.5}\sqrt{n}\,\mathrm{polylog}(n,d)$ in the adversarial setting and $Md^2\,\mathrm{polylog}(n,d)$ in the stochastic setting The parameter $M \in [d^{-1/2}, d^{-1/4}]$ depends on the geometry of the constraint set and the desired computational properties.

The previous best known rates were $d^{10.5}\sqrt{n}$ for an efficient algorithm [1], and $d^{2.5}\sqrt{n}$ for an inefficient algorithm using information-theoretic tools [2].

Our algorithm is built out of the following components:

1. The algorithm itself is adapted from [3] for the unconstrained case.
2. We introduce a bandit version of the extension proposed by [4].
3. To deal with the Adversarial case, negative bonuses, which can be seen as an increasing learning rate, are added.
4. A restart condition inspired by [1, 5] is added, which checks if the optimum is moving away and which guarantees negative regret on a restart.

Bandit Convex Optimisation

In rounds $t = 1, \dots, n$:

1. Learner chooses action $A_t \in K$
2. Learner suffers loss $\ell_t(A_t)$
3. Learner observes $Y_t = \ell_t(A_t) + \varepsilon_t$

Aim is to minimize the **Regret** with high probability:

$$\mathrm{Reg}_n = \sum_{t=1}^{n} \ell_t(A_t) - \min_{x \in K} \sum_{t=1}^{n} \ell_t(x).$$

Assumptions

Let

- $K \subset \mathbb{R}^d$, a convex body with
$$\mathbb{B}(1) \subseteq K \subseteq 2\mathbb{B}(d+1); \tag{1}$$
- Convex $\ell_1, \dots, \ell_n \colon K \to [0,1]$;
- In the **stochastic** case: $\ell = \ell_1 = \dots = \ell_n$;
- The noise, ε_t, is conditionally sub-Gaussian and mean 0.

Theoretical Results

Theorem 3. *There exists an algorithm such that with probability at least $1-\delta$,*

$$\mathrm{Reg}_n \lesssim \begin{cases} d^{3.5}\sqrt{n}\,\mathrm{polylog}(n,d,1/\delta) & \textit{Adversarial;} \\ d^{1.5}\sqrt{n}\,\mathrm{polylog}(n,d,1/\delta) & \textit{Stochastic, } K \textit{ in John's position or symmetric isotropic;} \\ d^{1.75}\sqrt{n}\,\mathrm{polylog}(n,d,1/\delta) & \textit{Stochastic, } K \textit{ isotropic.} \end{cases}$$

The parameters that achieve these regret results are

$$\gamma = \frac{1}{4dL} \qquad \eta = \sqrt{\frac{d}{nL^3}} \qquad \lambda = \frac{1}{d^3L^5} \qquad \sigma^2 = \frac{1}{d^2} \qquad \varepsilon = \frac{d^{3.5}L^{8.5}}{\sqrt{n}} \qquad \mathrm{F}_{\max} = d^5L^8\,.$$

In the stochastic setting we choose

$$\gamma = 0 \qquad \eta = \frac{Md}{\sqrt{n}} \qquad \lambda = \frac{5}{Md^{3/2}L^3} \qquad \sigma^2 = \frac{1}{16M^2dL^3} \qquad \varepsilon = \frac{Md^2L^5}{\sqrt{n}} \qquad \mathrm{F}_{\max} = 25M^2d^3L^5\,.$$

Convex Extension, Bandit edition

Let $\pi(x) = \inf\{t > 0 \colon x \in tK\}$, the **Minkowski functional**

Shrink $K \to K_\varepsilon = \{x \in \mathbb{R}^d \colon \pi_\varepsilon(x) = \frac{\pi(x)}{1-\epsilon} \le 1\}$

Define $\pi_+(x) = \max(1, \pi_\varepsilon(x))$, with $\varepsilon = \Theta(1/\sqrt{n})$

Definition 1 (Extension). *Let ℓ be a convex function, define the extension $f \colon \mathbb{R}^d \to \mathbb{R}$:*

$$f(x) = \pi_+(x)\ell\left(\frac{x}{\pi_+(x)}\right) + \frac{2(\pi_+(x)-1)}{\varepsilon} = \pi_+(x)\ell\left(\frac{x}{\pi_+(x)}\right) + 2v(x).$$

The output can be **estimated with only one query** X, set $A = X/\pi_+(X)$,

$$Y = \pi_+(X)[\ell(A) + \varepsilon] + 2v(X).$$

Notice that f is $\Theta(\sqrt{n})$-**Lipschitz**, which renders existing analyses vacuous. However, in our case we can still prove that $\sum_{t=1}^{n} Y_t^2 = \tilde{O}(n)$

Algorithm

1. **input** n, η, λ, γ, σ and $K_0 = K_\varepsilon$
2. **for** $t = 1$ **to** n
3. let $\Phi_{t-1}(x) = \frac{1}{2}\|x\|^2 + \sum_{u=1}^{t-1} \flat_u(x) + \eta \sum_{u=1}^{t-1} \hat{q}_u(x)$
4. compute $\mu_t = \arg\min_{x \in K_{t-1}} \Phi_{t-1}(x)$ and $\Sigma_t^{-1} = \Phi''_{t-1}(\mu_t)$
5. sample $X_t \sim \mathcal{N}(\mu_t, \Sigma_t)$
6. play $A_t = \frac{X_t}{\pi_+(X_t)}$ and observe $Y_t = \pi_+(X_t)[\ell_t(A_t) + \varepsilon_t] + 2v(X_t)$
7. $K_t = K_{t-1} \cap \{x : \|x - \mu_t\|_t^2 \le \mathrm{F}_{\max}\}$
8. **if** in the adversarial setting:
9. compute $z_t = \arg\min_{z \in \mathbb{R}^d} \sum_{s=1}^{t-1} \mathbf{1}(\flat_s \neq \mathbf{0}) \|z - \mu_s\|_s^2$
10. $$\flat_t(x) = \begin{cases} 0 & \text{if } \sum_{s=1}^{t-1} \mathbf{1}(\flat_s \neq \mathbf{0}) \|z_t - \mu_s\|_s^2 \ge \frac{\mathrm{F}_{\max}}{16} \\ -\gamma\|x - \mu_t\|_t^2 & \text{if } \|\cdot\|_t^2 \not\preceq \sum_{s=1}^{t-1} \mathbf{1}(\flat_s \neq \mathbf{0}) \|\cdot\|_s^2 \\ -\gamma\|x - \mu_t\|_t^2 & \text{if } \|\mu_t - z_t\|_t^2 \ge \frac{\mathrm{F}_{\max}}{8} \\ 0 & \text{otherwise.} \end{cases}$$
11. **if** $\max_{y \in K_t} \eta \sum_{u=1}^{t} (\hat{s}_u(\mu_u) - \hat{s}_u(y)) \le -\frac{\gamma \mathrm{F}_{\max}}{32}$
12. **then** restart algorithm
13. **end if**
14. **end for**

FTRL on quadratic surrogate estimates

Sample action
Observe extended loss

Update focus region

Determine
negative bonus

Check restart

Gaussian Smoothing

Definition 2 (Gaussian Smoothing). *Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be a convex function, and $X \sim \mathcal{N}(\mu, \Sigma)$, given parameter $\lambda \in \left(0, \frac{1}{1+d}\right)$, define*

$$s(z) = \mathbb{E}\left[\left(1 - \tfrac{1}{\lambda}\right) f(X) + \tfrac{1}{\lambda} f((1-\lambda)X + \lambda z)\right].$$

A **quadratic approximation** is defined as

$$q(z) = \langle s'(\mu), z - \mu \rangle + \tfrac{1}{4}\|z - \mu\|_{s''(\mu)}^2\,.$$

All necessary quantities can be estimated. Let $r(X,z) = \frac{p\left(\frac{X - \lambda z}{1-\lambda}\right)}{(1-\lambda)^d p(X)}$ with p the density of the $\mathcal{N}(\mu, \Sigma)$-distribution. The relevant estimators are

- $\hat{s}(z) = Y\left(1 + \dfrac{r(X,z) - 1}{\lambda}\right)$
- $\hat{s}'(\mu) = \dfrac{Y r(X,\mu)}{1-\lambda} \Sigma^{-1} \left(\dfrac{X - \mu}{1-\lambda}\right)$
- $\hat{s}''(\mu) = \frac{\lambda Y r(X,\mu)}{(1-\lambda)} \left(\Sigma^{-1}\left[\frac{X-\mu}{1-\lambda}\right]\left[\frac{X-\mu}{1-\lambda}\right]^\top \Sigma^{-1} - \Sigma^{-1}\right)$
- $\hat{q}(z) = \langle \hat{s}'(\mu), z - \mu \rangle + \frac{1}{4}\|z - \mu\|_{\hat{s}''(\mu)}^2$

Only concentrate well in the focus regions
$\{x \in K \colon \lambda \|z - \mu_t\|_{\Sigma_t^{-1}} \le \frac{1}{L}\}$

Geometry of the Constraint Set

The dimension dependence in the stochastic case of our algorithm depends on **the mean width of the polar body** K° of the constraints set K:

$$M(K^\circ) = \int_{\mathbb{S}(1)} \pi(x)\,\mathrm{d}\rho(x).$$

The parameter of interest is $M = \max\left(d^{-1/2}, M(K^\circ)\right)$ and dimension dependence in the stochastic case is controlled by Md^2.

(a) Without any assumption on K you can take $M = d^{-1/2}$, but the algorithm may be computationally **inefficient**.

(b) Given access to sampling and membership oracles for K you can take $M = d^{-1/4}$ and the algorithm is **efficient**.

(c) Given access to sampling and membership oracles for a symmetric K you can take $M = d^{-1/2}$ and the algorithm is **efficient**.

References & Link

[1] S. Bubeck, Y.T. Lee, and R. Eldan. Kernel-based methods for bandit convex optimization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 72–85, New York, NY, USA, 2017. ACM.

[2] T. Lattimore. Improved regret for zeroth-order adversarial bandit convex optimisation. *Mathematical Statistics and Learning*, 2(3/4):311–334, 2020.

[3] T. Lattimore and A. György. A second-order method for stochastic bandit convex optimisation. In *Conference on Learning Theory*, 2023.

[4] Z. Mhammedi. Efficient projection-free online convex optimization with membership oracle. In *Conference on Learning Theory*, pages 5314–5390, 2022.

[5] A. Suggala, P. Ravikumar, and P. Netrapalli. Efficient bandit convex optimization: Beyond linear losses. In *Conference on Learning Theory*, pages 4008–4067, 2021.

Read our paper online!