Open Problem: Regret Bounds for Noise-Free Kernel-Based Bandits

Sattar Vakili SATTAR.VAKILI@MTKRESEARCH.COM
MediaTek Research

Editors: Po-Ling Loh and Maxim Raginsky

Abstract

Kernel-based bandit is an extensively studied black-box optimization problem, in which the objective function is assumed to live in a known reproducing kernel Hilbert space. While nearly optimal regret bounds (up to logarithmic factors) are established in the noisy setting, surprisingly, less is known about the noise-free setting (when the exact values of the underlying function is accessible without observation noise). We discuss several upper bounds on regret; none of which seem order optimal, and provide a conjecture on the order optimal regret bound.

Keywords: Kernel-based bandit, Bayesian optimization, Gaussian processes

1. Introduction

Black-box optimization of a (possibly non-convex) function from expensive evaluations is a ubiquitous problem in machine learning, including both academic research and industrial applications. Those include A/B testing, hyperparameter tuning (including AlphaGo, Chen et al., 2018), robotics, environmental monitoring, and more (Shahriari et al., 2016). A particularly successful approach is based on the use of Gaussian process modeling, due to its versatility, modeling power, and ability to provide uncertainty estimates. A mathematical formulation of the problem under the bandit setting has been extensively studied in the literature. Nearly optimal regret bounds are established in the noisy setting. Surprisingly, the problem of order optimal regret bounds remains open under the noise-free setting. Here, we overview the existing results and give a formal description of the open problem, aspiring to motivate solutions.

2. Problem Setup

Consider the sequential optimization of an objective function $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \subset \mathbb{R}^d$ is a compact set. A learning algorithm is allowed to collect a sequence of observations $\{(x_i, y_i)\}_{i=1}^{\infty}$, where $y_i = f(x_i)$ in the noise-free setting, and $y_i = f(x_i) + \epsilon_i$ in the noisy setting with ϵ_i being a well behaved observation noise. The objective is to get as close as possible to the maximum of f. The performance of the algorithm is measured in terms of (cumulative) regret, defined as the cumulative loss in the values of the objective function at observation points, compared to a global maximum:

$$\mathcal{R}(N) = \sum_{i=1}^{N} \left(f(x^*) - f(x_i)\right), \tag{1}$$

where $x^* \in \mathrm{argmax}_{x\in\mathcal{X}} f(x)$ is a global maximum. The regularity assumption on f which makes the problem tractable is given next.

The RKHS and the regularity assumption on f: Consider a positive definite kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ with respect to a finite Borel measure. Let $\mathcal{H}_k$ denote the reproducing kernel Hilbert space (RKHS) corresponding to k, defined as a Hilbert space equipped with an inner product $\langle ., . \rangle_{\mathcal{H}_k}$ satisfying the following: $k(., x) \in \mathcal{H}_k$, $\forall x \in \mathcal{X}$, and $\langle f, k(., x) \rangle_{\mathcal{H}_k} = f(x)$, $\forall x \in \mathcal{X}, \forall f \in \mathcal{H}_k$ (reproducing property). The typical assumption in kernel-based models is that the objective function f satisfies $f \in \mathcal{H}_k$ for a known kernel k.

Assumption 1 *Assume f is fixed with $\|f\|_{\mathcal{H}_k} \leq C_k$, for some $C_k > 0$, where $\|f\|_{\mathcal{H}_k} = \sqrt{\langle f, f \rangle_{\mathcal{H}_k}}$.*

Under Assumption 1, the sequential optimization problem given above is often referred to as that of kernel-based (kernelized) bandit, Gaussian process (GP) bandit, or Bayesian optimization. The latter two terms are motivated by the algorithm design which often employs a GP surrogate model.

GP surrogate model: It is useful for algorithm design to employ a centered GP model F with kernel k, which provides a surrogate posterior mean (prediction) and a surrogate posterior variance (uncertainty estimate) for the kernel-based model. Defining $\mu_n(x) = \mathbb{E}\big[F(x) | \{(x_i, y_i)\}_{i=1}^n\big]$ and $\sigma_n^2(x) = \mathbb{E}\big[(F(x) - \mu_n(x))^2 | \{(x_i, y_i)\}_{i=1}^n\big]$, it is well known that $\mu_n(x) = \mathbf{k}_n^\top(x)(\lambda^2 \mathbf{I}_n + \mathbf{K}_n)^{-1}\mathbf{y}_n$ and $\sigma_n^2(x) = k(x, x) - \mathbf{k}_n^\top(x)(\lambda^2 \mathbf{I}_n + \mathbf{K}_n)^{-1}\mathbf{k}_n(x)$, where $\mathbf{y}_n = [y_1, \ldots, y_n]^\top$, $\mathbf{k}_n(x) = [k(x, x_1), \ldots, k(x, x_n)]^\top$, $\mathbf{K}_n$ is the positive definite kernel matrix $[\mathbf{K}_n]_{i,j} = k(x_i, x_j)$, $\mathbf{I}_n$ is the identity matrix of dimension n, and λ^2 is the variance of a zero mean Gaussian surrogate distribution for the noise. In the noise-free setting, one may set $\lambda = 0$ in the above expressions.

The Matérn family of kernels is perhaps the most commonly used in practice (see, e.g., Snoek et al., 2012; Shahriari et al., 2016) and theoretically interesting (see, e.g., Srinivas et al., 2010; Bull, 2011) family of kernels. It is known that the RKHS corresponding to a Matérn kernel is equivalent to a Sobolev space (see, e.g., Teckentrup, 2018) that provides an intuitive interpretation of the function class based on its smoothness. Furthermore, the Matérn kernels may be insightful for the theory of neural networks, as it has been shown that the neural tangent kernel (Jacot et al., 2018) is equivalent to a Matérn kernel (Vakili et al., 2021a). Following the literature, we also emphasize the Matérn family in the formulation of the open problem.

3. Overview of the Results in the Noisy Setting

We first overview the main results for the kernel-based bandit problem in the noisy setting and under Assumption 1. Classical algorithms such as GP upper confidence bound (GP-UCB, Srinivas et al., 2010; Chowdhury and Gopalan, 2017), Thompson sampling (GP-TS, Chowdhury and Gopalan, 2017), and expected improvement (EI, Gupta et al., 2022) sequentially select the observation points based on a score referred to as acquisition function. The best known regret bounds for these algorithms scales as $\tilde{\mathcal{O}}(\Gamma_{k,\lambda}(N)\sqrt{N})$ over N steps (see the references above), where $\Gamma_{k,\lambda}(N)$ is a kernel specific complexity term, referred to as the *maximal information gain* between the noisy observations and the latent GP surrogate model (Srinivas et al., 2010):

$$\Gamma_{k,\lambda}(n) = \sup_{\{x_i\}_{i=1}^n \subset \mathcal{X}} \frac{1}{2} \log \det(\mathbf{I}_n + \frac{1}{\lambda^2}\mathbf{K}_n). \tag{2}$$

That is also nearly equivalent to the *effective dimension* of the kernel for a dataset of finite size n (Calandriello et al., 2019). Kernel specific bounds on $\Gamma_{k,\lambda}(n)$, with a focus on Matérn family, are provided in Srinivas et al. (2010); Vakili et al. (2021b,a); Kassraie and Krause (2022).

The $\tilde{\mathcal{O}}(\Gamma_{k,\lambda}(N)\sqrt{N})$[1] scaling is not tight in general, and may even fail to be sublinear in many cases of interest, since $\Gamma_{k,\lambda}(N)$ may grow faster than $\sqrt{N}$. It remains an open problem whether the suboptimal regret bounds of these acquisition based algorithms is a fundamental limitation or a shortcoming of their proof (see Vakili et al., 2021c, for the details).

On discrete domains, the SupKernelUCB algorithm was shown to have an $\tilde{\mathcal{O}}(\sqrt{\Gamma_{k,\lambda}(N)N})$ regret bound (Valko et al., 2013). SupKernelUCB is not considered to be practical. Several more practical algorithms (which also apply to continuous domains) with $\tilde{\mathcal{O}}(\sqrt{\Gamma_{k,\lambda}(N)N})$ regret have been proposed recently: a tree-based domain-shrinking algorithm (GP-ThreDS, Salgia et al., 2021), Robust Inverse Propensity Score for experimental design (RIPS, Camilleri et al., 2021), batched pure exploration (BPE, Li and Scarlett, 2022), and its sparse variation (S-BPE, Vakili et al., 2022).

Substituting the bounds on $\Gamma_{k,\lambda}(N)$ for the Matérn kernels leads to $\mathcal{R}(N) = \tilde{\mathcal{O}}(N^{\frac{\nu+d}{2\nu+d}})$, where ν is the smoothness parameter of the kernel. Scarlett et al. (2017) proved a matching (up to logarithmic factors) $\Omega(N^{\frac{\nu+d}{2\nu+d}})$ lower bound on the regret, which shows the order optimality of the $\tilde{\mathcal{O}}(\sqrt{\Gamma_{k,\lambda}(N)N})$ regret bounds in this case.

4. Regret Bounds in the Noise-Free Setting

As outlined above, there are many results on the kernel-based bandit problem under the noisy setting, several of which achieving nearly optimal regret bounds. Under the noise-free setting, Lyu et al. (2019) is the only work providing bounds on the cumulative regret, while Bull (2011) provided nearly optimal bounds on the simple regret. In this section, we first overview these two works, and then provide a formal description of the open problem of cumulative regret under the noise-free setting.

4.1. From Noisy to Noise-Free

One straightforward way to establish regret bounds under the noise-free setting is to set the noise equal to zero in the results under the noisy setting. Lyu et al. (2019) took this approach and showed an $\mathcal{O}(\sqrt{\Gamma_{k,\lambda}(N)N})$ bound on the cumulative regret for GP-UCB. The analysis mainly follows that of GP-UCB under the noisy setting (Srinivas et al., 2010), except for using tighter confidence intervals for GP models under the noise-free setting. In the case of Matérn kernels, this implies an $\mathcal{R}(N) = \tilde{\mathcal{O}}(N^{\frac{\nu+d}{2\nu+d}})$ regret that always grows faster than $\sqrt{N}$, even for highly smooth kernels (large ν). As we will see next, this regret bound is not order optimal in the noise-free setting.

4.2. Explore then Commit

Consider the simple regret variation of the kernel-based bandit problem. In this variation, after n (possibly adaptive) observations $\{(x_i, y_i)\}_{i=1}^n$, the algorithm selects a candidate maximizer $\hat{x}_n$, and its performance is measured in terms of simple regret defined as $r_n = f(x^*) - f(\hat{x}_n)$. Focusing

1. The notations $\mathcal{O}$ and $\tilde{\mathcal{O}}$ are used for mathematical order, and that up to hiding logarithmic factors, respectively.

on Matérn family and simple regret, (Bull, 2011) proved an $r_n = \tilde{\mathcal{O}}(n^{-(\min\{\nu,1\})/d})$ regret bound for the EI algorithm (see their Theorem 2), that translates to an $R(N) = \mathcal{O}(N^{(d-1)/d})$ bound on its cumulative regret when $\nu \geq 1$ (which is typically the case, except for the Laplace kernel where $\nu = \frac{1}{2}$). Furthermore, they showed that EI mixed with pure exploration achieves an $r_n = \tilde{\mathcal{O}}(n^{-\frac{\nu}{d}})$ simple regret (see their Theorem 5), which cannot be improved (see their Theorem 1). The cumulative regret $R(N)$ for this algorithm, however, grows linearly in N, as a consequence of pure exploration. The question of cumulative regret appears more challenging due to the exploration-exploitation tradeoff, which is inherent to the bandit problems.

Inspired by the results of Bull (2011), and using the *explore then commit* technique (e.g., see Garivier et al., 2016), we can design a simple algorithm with sublinear regret. Specifically, consider an algorithm which performs pure exploration by, e.g., choosing nearly uniform points across the domain, up to step N_0. The algorithm then commits to the best observation point based on the GP prediction: $\hat{x}_{N_0} = \arg\max_{x\in\mathcal{X}} \mu_{N_0}(x)$ (with $\lambda = 0$ in the expression of μ_{N_0}), for the remaining of the steps: $x_i = \hat{x}_{N_0}, \forall i > N_0$. Choosing the optimum value for N_0, leads to a cumulative regret of $\mathcal{R}(N) = \mathcal{O}(N^{\frac{d}{\nu+d}})$. Note that this regret bound is tighter than the one in Lyu et al. (2019) given above, indicating that the analytical techniques used in the noisy setting may not be suitable for the noise-free setting.

4.3. Open Problem

Problem 1 Consider the kernel-based bandit problem given in Section 2 under the noise-free setting ($\epsilon_i = 0, \forall i$). What is the lowest growth rate of $R(N)$ with N? As a specific case of interest, when the kernel k is a Matérn kernel with smoothness parameter ν and $R(N) = \tilde{\mathcal{O}}(N^{\alpha})$, what is the smallest value of α, achievable by a learning algorithm?

4.4. Discussion

In the case of Matérn kernels, as outlined above, the existing upper bounds include $\tilde{\mathcal{O}}(N^{\frac{\nu+d}{2\nu+d}})$ (Lyu et al., 2019), $\mathcal{O}(N^{\frac{d-1}{d}})$ (EI, Bull, 2011), and $\tilde{\mathcal{O}}(N^{\frac{d}{\nu+d}})$ (explore then commit). We however conjecture that the following regret bound is achievable:

$$\mathcal{R}(N) = \begin{cases} \mathcal{O}(N^{(d-\nu)/d}), & \text{when } d > \nu, \\ \mathcal{O}\left(\log(N)\right), & \text{when } d = \nu, \\ \mathcal{O}(1), & \text{when } d < \nu. \end{cases} \tag{3}$$

We here give an informal reasoning for this claim. Following standard UCB-based bandit algorithm techniques, it can be shown that the GP-UCB algorithm attains $\mathcal{R}(N) = \mathcal{O}(C_k \sum_{i=1}^{N} \sigma_{i-1}(x_i))$. Let $\{x_i^*\}_{i=1}^n \in \arg\max_{\{x_i\}_{i=1}^n \subset \mathcal{X}} \sum_{i=1}^n \sigma_{i-1}(x_i)$ and $\Theta_n^* = \sum_{i=1}^n \sigma_{i-1}(x_i^*)$. We then have $\mathcal{R}(N) = \mathcal{O}(C_k \Theta_N^*)$, which reduces the problem to bounding Θ_n^*. We conjecture that under certain mild regularity of the domain, $\{x_i^*\}_{i=1}^n$ are distributed nearly uniformly across the domain, in which case $\sigma_{i-1}(x_i^*) = \mathcal{O}(i^{-\frac{\nu}{d}})$ (see, e.g., Teckentrup, 2018, Lemma 3.8). That, when summed over i, leads to $\Theta_n^* = \mathcal{O}(n^{\frac{d-\nu}{d}})$ when $d > \nu$, $\Theta_n^* = \mathcal{O}\left(\log(n)\right)$ when $d = \nu$, and $\Theta_n^* = \mathcal{O}(1)$ when $d < \nu$.

References

Adam D Bull. Convergence rates of efficient global optimization algorithms. *Journal of Machine Learning Research*, 12(10), 2011.

Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, and Lorenzo Rosasco. Gaussian process optimization with adaptive sketching: Scalable and no regret. In *Conference on Learning Theory*, 2019.

Romain Camilleri, Kevin Jamieson, and Julian Katz-Samuels. High-dimensional experimental design and kernel bandits. In *International Conference on Machine Learning*, pages 1227–1237. PMLR, 2021.

Yutian Chen, Aja Huang, Ziyu Wang, Ioannis Antonoglou, Julian Schrittwieser, David Silver, and Nando de Freitas. Bayesian optimization in AlphaGo. *arXiv preprint arXiv:1812.06855*, 2018.

Sayak Ray Chowdhury and Aditya Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853, 2017.

Aurélien Garivier, Tor Lattimore, and Emilie Kaufmann. On explore-then-commit strategies. *Advances in Neural Information Processing Systems*, 29, 2016.

Sunil Gupta, Santu Rana, Svetha Venkatesh, et al. Regret bounds for expected improvement algorithms in gaussian process bandit optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 8715–8737. PMLR, 2022.

Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. *arXiv preprint arXiv:1806.07572*, 2018.

Parnian Kassraie and Andreas Krause. Neural contextual bandits without regret. In *International Conference on Artificial Intelligence and Statistics*, 2022.

Zihan Li and Jonathan Scarlett. Gaussian process bandit optimization with few batches. In *International Conference on Artificial Intelligence and Statistics*, 2022.

Yueming Lyu, Yuan Yuan, and Ivor W Tsang. Efficient batch black-box optimization with deterministic regret bounds. *arXiv preprint arXiv:1905.10041*, 2019.

Sudeep Salgia, Sattar Vakili, and Qing Zhao. A domain-shrinking based Bayesian optimization algorithm with order-optimal regret performance. *Conference on Neural Information Processing Systems*, 34, 2021.

Jonathan Scarlett, Ilija Bogunovic, and Volkan Cevher. Lower bounds on regret for noisy Gaussian process bandit optimization. In *Conference on Learning Theory*, pages 1723–1742, 2017.

B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.

Niranjan Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *International Conference on Machine Learning*, pages 1015–1022. Omnipress, 2010.

Aretha L. Teckentrup. Convergence of gaussian process regression with estimated hyper-parameters and applications in bayesian inverse problems. *Available at Arxiv.*, 2018.

Sattar Vakili, Michael Bromberg, Jezabel Garcia, Da-shan Shiu, and Alberto Bernacchia. Uniform generalization bounds for overparameterized neural networks. *arXiv preprint arXiv:2109.06099*, 2021a.

Sattar Vakili, Kia Khezeli, and Victor Picheny. On information gain and regret bounds in Gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90, 2021b.

Sattar Vakili, Jonathan Scarlett, and Tara Javidi. Open problem: Tight online confidence intervals for RKHS elements. In *Conference on Learning Theory*, pages 4647–4652. PMLR, 2021c.

Sattar Vakili, Jonathan Scarlett, Da-shan Shiu, and Alberto Bernacchia. Improved convergence rates for sparse approximation methods in kernel-based learning. In *International Conference on Machine Learning*, 2022.

Michal Valko, Nathan Korda, Rémi Munos, Ilias Flaounas, and Nello Cristianini. Finite-time analysis of kernelised contextual bandits. In *Conference on Uncertainty in Artificial Intelligence*, page 654–663, 2013.