Optimal Order Simple Regret for Gaussian Process Bandits

Satar Vakili, *Nacime Bouziani, Sepehr Jalali,*
Alberto Bernacchia, Da-shan Shiu

MediaTek Research

NeurIPS 2021

Overview

- $\diamond$ Simple regret in Gaussian process bandit problem (kernel-based bandit, Bayesian optimization)
- $\diamond$ We prove an $\tilde{\mathcal{O}}(\sqrt{\frac{\gamma_N}{N}})$ simple regret
- $\diamond$ That is tight up to logarithmic factors
- $\diamond$ We formalize confidence intervals for RKHS elements which may be of broader interest

Problem Formulation: Setting

Zeroth-order optimization

- ◇ Consider optimization of an objective function $f : \mathcal{X} \to \mathbb{R}$, $\mathcal{X} \subset \mathbb{R}^d$
- ◇ From a sequence of n noisy observations $\{(x_i, y_i)\}_{i=1}^n$, $y_i = f(x_i) + \epsilon_i$
- ◇ "Zeroth-order" signifies direct observations from f, and not from its gradient for example
- ◇ $x^* \in \arg\max_{x \in \mathcal{X}} f(x)$
- ◇ Objective: to get as close as possible to $f(x^*)$,

Problem Formulation: Algorithm

Algorithm $\mathcal{A}$

- ◇ Consider an algorithm $\mathcal{A}$
- ◇ Selects a sequence of observation points $\{x_i\}_{i=1}^n$
- ◇ Receives noisy observations $\{y_i = f(x_i) + \epsilon_i\}_{i=1}^n$
- ◇ Predicts a candidate maximizer $\hat{x}_n^*$
- ◇ Performance of $\mathcal{A}$ is measured by simple regret

$$r_n^{\mathcal{A}} = f(x^*) - f(\hat{x}_n^*)$$

Problem Formulation: Regularity Assumptions

◇ Assumption 1: $f \in \mathcal{H}_k$, the reproducing kernel Hilbert space (RKHS) corresponding to a positive definite kernel k

$$\|f\|_{\mathcal{H}_k} \leq B$$

◇ Assumption 2: Sub-Gaussian noise with parameter R

$$\forall h \in \mathbb{R}, \forall n \in \mathbb{N}, \mathbb{E}[e^{h\epsilon_n}] \leq \exp(\frac{h^2 R^2}{2}),$$

◇ Assumption 3 : Light-tailed noise with parameters h_0, ξ_0

$$\forall h \leq h_0, \forall n \in \mathbb{N}, \mathbb{E}[e^{h\epsilon_n}] \leq \exp(\frac{h^2 \xi_0}{2}), \text{ for some } \xi_0 > 0$$

▶ We solve the problem under Assumption 1 **and** (2 **or** 3)

Surrogate Gaussian Process Model

Mean

$$\mu_n(x) = k^\top(x, X_n)\left(k(X_n, X_n) + \lambda^2 I_n\right)^{-1} Y_n$$

Variance

$$k_n(x, x') = k(x, x') - k^\top(x, X_n)\left(k(X_n, X_n) + \lambda^2 I_n\right)^{-1} k(x', X_n), \;\; \sigma_n^2(x) = k_n(x, x)$$

- $\diamond$ $k(x, X_n) = [k(x, x_1), k(x, x_2), \ldots, k(x, x_n)]^\intercal$
- $\diamond$ $k(X_n, X_n) = [k(x_i, x_j)]_{i,j=1}^n$ is the covariance matrix
- $\diamond$ $\lambda > 0$ is a real number

Confidence Intervals

Confidence intervals for RKHS elements with sub-Gaussian noise

◇ Provided n observations $\{X_n, Y_n\}$

◇ X_n independent of E_n

◇ For a fixed $x \in \mathcal{X}$,

$$f(x) \leq \mu_n(x) + (B + \beta(\delta))\sigma_n(x), \text{ with probability } 1 - \delta$$

$$f(x) \geq \mu_n(x) - (B + \beta(\delta))\sigma_n(x), \text{ with probability } 1 - \delta$$

$$\beta(\delta) = \frac{R}{\lambda}\sqrt{2\log(\frac{1}{\delta})}$$

Confidence Intervals

Confidence intervals for RKHS elements with light-tailed noise

◇ Provided n observations $\{X_n, Y_n\}$

◇ X_n independent of E_n

◇ For a fixed $x \in \mathcal{X}$,

$$f(x) \leq \mu_n(x) + (B + \beta(\delta))\sigma_n(x), \text{ with probability } 1 - \delta$$

$$f(x) \geq \mu_n(x) - (B + \beta(\delta))\sigma_n(x), \text{ with probability } 1 - \delta$$

$$\beta(\delta) = \frac{1}{\lambda}\sqrt{2\left(\xi_0 \vee \frac{2\log(\frac{1}{\delta})}{h_0^2}\right)\log(\frac{1}{\delta})}$$

Posterior Variance of the GP Model

Proposition: For the posterior variance of the surrogate GP model, we have

$$\sigma_n^2(x) = \sup_{f:||f||_{\mathcal{H}_k} \le 1} (f(x) - Z_n^\top(x) F_n)^2 + \lambda^2 \|Z_n(x)\|_{l^2}^2.$$

◇ $Z_n^\top(x) = k^\top(x, X_n) \left(k(X_n, X_n) + \lambda^2 I_n \right)^{-1}$.

◇ maximum prediction error for RKHS elements

◇ noise variance

Pure Exploration Algorithm

Maximum variance reduction

Algorithm 1 Maximum Variance Reduction (MVR)

1: **Initialization:** k, $\mathcal{X}$, f, $\sigma_0^2(x) = k(x, x)$.
2: **for** $n = 1, 2, \ldots, N$ **do**
3: $\quad x_n = \text{argmax}_{x \in \mathcal{X}} \sigma_{n-1}^2(x)$, where a tie is broken arbitrarily.
4: $\quad$ Update $\sigma_n^2(.)$ according to (2).
5: **end for**
6: Update $\mu_N(.)$ according to (2).
7: **return** $\hat{x}_N^* = \text{argmax}_{x \in \mathcal{X}} \mu_N(x)$, where a tie is broken arbitrarily.

$$x_n = \arg\max_{x \in \mathcal{X}} \sigma_{n-1}^2(x)$$

Information Gain

◇ Mutual information between Y_n and f

$$\mathcal{I}(Y_n; f) = \log\det\left(I_n + \frac{1}{\lambda^2}k(X_n, X_n)\right)$$

◇ $\mathcal{I}(Y_n; f) \sim$ effective dimension of the kernel

◇ Maximal information gain γ_N

$$\gamma_N = \sup_{X_n \subset \mathcal{X}} \mathcal{I}(Y_n; f)$$

◇ Matérn: $\gamma_N = \mathcal{O}\left(N^{\frac{d}{2\nu+d}}(\log(N))^{\frac{2\nu}{2\nu+d}}\right)$,
Squared Exponential: $\gamma_N = \mathcal{O}\left((\log(N))^{d+1}\right)$ [Srinivas et al., 2010, Vakili et al., 2020]

Simple Regret of MVR

Theorem: Under Assumption 1 and (2 or 3), with probability at least $1-\delta$

$$r_N^{\mathsf{MVR}} \leq \sqrt{\frac{2\gamma_N}{\log(1+\frac{1}{\lambda^2})N}} \left(2B + \beta(\tfrac{\delta}{3}) + \beta\left(\frac{\delta}{3C\big(B+\sqrt{N}\beta(2\delta/3N)\big)^d N^{d/2}}\right)\right) + \frac{2}{\sqrt{N}}.$$

◇ Assumption 2: $\beta(\delta) = \frac{R}{\lambda}\sqrt{2\log(\frac{1}{\delta})}$

◇ Assumption 3: $\beta(\delta) = \frac{1}{\lambda}\sqrt{2\left(\xi_0 \vee \frac{2\log(\frac{1}{\delta})}{h_0^2}\right)\log(\frac{1}{\delta})}$

Simple Regret of MVR

◇ Assumption 2:

$$r_N^{\mathsf{MVR}} = \mathcal{O}(\sqrt{\frac{\gamma_N \log(N^d/\delta)}{N}})$$

◇ Assumption 3:

$$r_N^{\mathsf{MVR}} = \mathcal{O}\left(\sqrt{\frac{\gamma_N}{N}} \log(N^d/\delta)\right)$$

◇ In the case of Matérn ν

$$r_N^{\mathsf{MVR}} = \mathcal{O}\left(N^{\frac{-\nu}{2\nu+d}} (\log(N))^{\frac{\nu}{2\nu+d}} \sqrt{\log(N^d/\delta)}\right)$$

$$r_N^{\mathsf{MVR}} = \mathcal{O}\left(N^{\frac{-\nu}{2\nu+d}} (\log(N))^{\frac{\nu}{2\nu+d}} \log(N^d/\delta)\right)$$

Sample Complexity

◇ Define

$$N_\epsilon = \min\{N \in \mathbb{N} : \mathbb{E}[r_n^{\text{MVR}}] \leq \epsilon, \forall n \geq N\}$$

◇ As a result of simple regret we have

Kernel	Under Assumption 2	Under Assumption 3
SE	$N_\epsilon = \mathcal{O}\left((\frac{1}{\epsilon})^2 \log(\frac{1}{\epsilon})^{d+2}\right)$	$N_\epsilon = \mathcal{O}\left((\frac{1}{\epsilon})^2 \log(\frac{1}{\epsilon})^{d+3}\right)$
Matérn-ν	$N_\epsilon = \mathcal{O}\left((\frac{1}{\epsilon})^{2+\frac{d}{\nu}} (\log(\frac{1}{\epsilon})^{\frac{4\nu+d}{2\nu}})\right)$	$N_\epsilon = \mathcal{O}\left((\frac{1}{\epsilon})^{2+\frac{d}{\nu}} (\log(\frac{1}{\epsilon})^{\frac{6\nu+2d}{2\nu}})\right)$

◇ These bounds match the lower bounds given in Scarlett et al. [2017], up to log factors

Discussion and Open Problem

◇ Recall our confidence interval width multiplier $B + \frac{R}{\lambda}\sqrt{2\log(\frac{1}{\delta})}$

◇ Chowdhury and Gopalan [2017]: $B + R\sqrt{2(\gamma_n + 1 + \log(\frac{1}{\delta}))}$

◇ That results in $\mathcal{O}(\frac{\gamma_N}{\sqrt{N}})$ regret for typical algorithms such as GP-UCB and GP-TS (that is not always sublinear)

◇ Neither confidence intervals imply the other

◇ A tight analysis of GP-UCB and GP-TS in the RKHS setting remains an open problem [Vakili et al., 2021]

References

S. R. Chowdhury and A. Gopalan. On kernelized multi-armed bandits. In *International Conference on Machine Learning*, pages 844–853, 2017.

J. Scarlett, I. Bogunovic, and V. Cevher. Lower bounds on regret for noisy Gaussian process bandit optimization. In *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1723–1742, Amsterdam, Netherlands, 07–10 Jul 2017. PMLR.

N. Srinivas, A. Krause, S. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, pages 1015–1022. Omnipress, 2010.

S. Vakili, K. Khezeli, and V. Picheny. On information gain and regret bounds in gaussian process bandits. *arXiv preprint arXiv:2009.06966*, 2020.

S. Vakili, J. Scarlett, and T. Javidi. Open problem: Tight online confidence intervals for rkhs elements. In *Conference on Learning Theory*, pages 4647–4652. PMLR, 2021.