On Information Gain and Regret Bounds in Gaussian Process Bandits

Sattar Vakili [*], *Kia Khezeli* [†], *Victor Picheny* [‡]

[*]*MediaTek Research*, , [†]*Mayo Clinic*, [‡]*Secondmind Labs*

April 2021

Gaussian Process Bandit

- A decision domain $\mathcal{X}$ ($x \in \mathcal{X}$)
- An objective function f
- A learner or decision maker π selects x_t

$$x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow \ldots \rightarrow x_T$$

- Noisy observations

$$y_t = f(x_t) + \epsilon_t$$

- Objective: minimize **Regret**

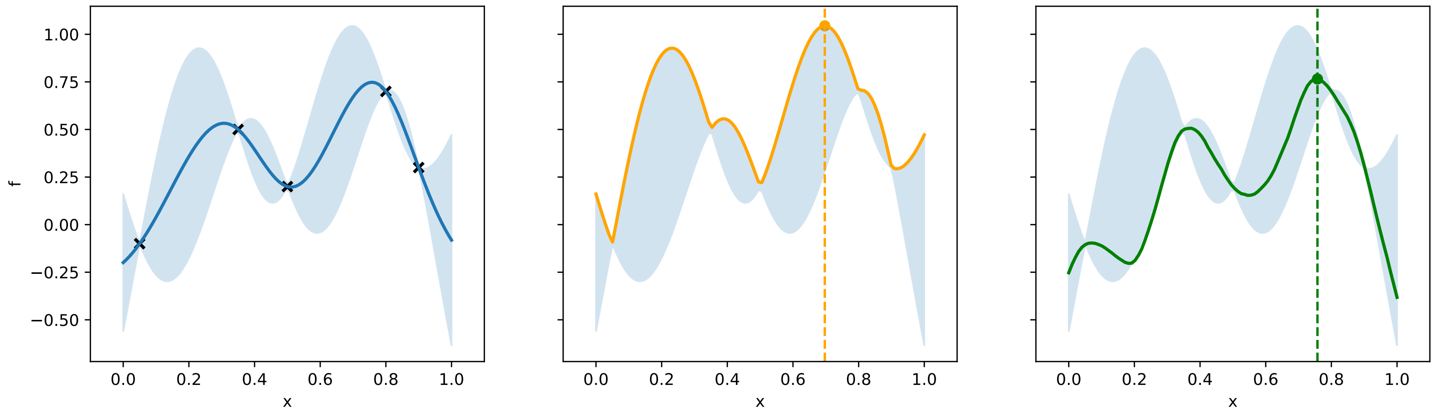
$$\textit{minimize } \sum_{t=1}^{T} f(x^*) - f(x_t)$$

Gaussian Process Bandit

- Bayesian Regularity Assumptions
  - f is a sample from a GP
  - Noise is Gaussian
- Frequentist Regularity Assumptions
  - f belongs to a Reproducing Kernel Hilbert Space (RKHS)
  - Noise is sub-Gaussian

Algorithms

- GP-UCB (middle): optimistic optimization [Srinivas et al., 2010]
- GP-TS (right): posterior sampling [Chowdhury and Gopalan, 2017]



Regret Bounds and Information Gain

Regret Bounds for GP bandits (Informal):

$$R(T) = \mathcal{O}(\sqrt{T\gamma_T})$$

where γ_T denotes the maximal information gain

$$\gamma_T = \sup_{\mathbf{X}_T \subseteq \mathcal{X}} I(\mathbf{y}_T; \hat{f}),$$

$$I(\mathbf{y}_t; \hat{f}) = \frac{1}{2} \log \det(\mathbf{I}_t + \frac{1}{\tau} K_{\mathbf{X}_t, \mathbf{X}_t}).$$

Main Result: Theory

Theorem (Bound on γ_T): Consider a GP with a kernel k. For $D \in \mathbb{N}$ The following upper bound on γ_T holds for all $D \in \mathbb{N}$.

$$\gamma_T \le \frac{1}{2} D \log\left(1 + \frac{\bar{k}T}{\tau D}\right) + \frac{1}{2}\frac{\delta_D T}{\tau}.$$

The expression can be simplified as

$$\gamma_T = \mathcal{O}\left(D\log(T) + \delta_D T\right)$$

- $\delta_D = \sum_{m=D+1}^{\infty} \lambda_m \psi^2$

Main Results: Improved Bounds

Kernel	Bound on γ_T	Regret Lower Bound	Regret Upper Bound
Polynomial eigendecay	$\mathcal{O}\left(T^{\frac{1}{\beta_p}}\log^{1-\frac{1}{\beta_p}}(T)\right)$	-	$\tilde{\mathcal{O}}\left(T^{\frac{\beta_p+1}{2\beta_p}}\right)$
Exponential eigendecay	$\mathcal{O}\left(\log^{1+\frac{1}{\beta_e}}(T)\right)$	-	$\tilde{\mathcal{O}}\left(T^{\frac{1}{2}}\log^{\frac{1}{2\beta_e}}(T)\right)$
Matérn$-\nu$	$\mathcal{O}\left(T^{\frac{d}{2\nu+d}}\log^{\frac{2\nu}{2\nu+d}}(T)\right)$	$\Omega(T^{\frac{\nu+d}{2\nu+d}})$	$\tilde{\mathcal{O}}\left(T^{\frac{\nu+d}{2\nu+d}}\right)$
SE	$\mathcal{O}\left(\log^{d+1}(T)\right)$	$\Omega\left(T^{\frac{1}{2}}\log^{\frac{d}{2}}(T)\right)$	$\tilde{\mathcal{O}}\left(T^{\frac{1}{2}}\log^{\frac{d}{2}}(T)\right)$

- Polynomial eigendecay: $\lambda_m = \mathcal{O}(m^{-\beta_p})$, $\beta_p > 1$
- Exponential eigendecay: $\lambda_m = \mathcal{O}(\exp(-m^{\beta_e}))$, $\beta_e > 0$,

References

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

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.*