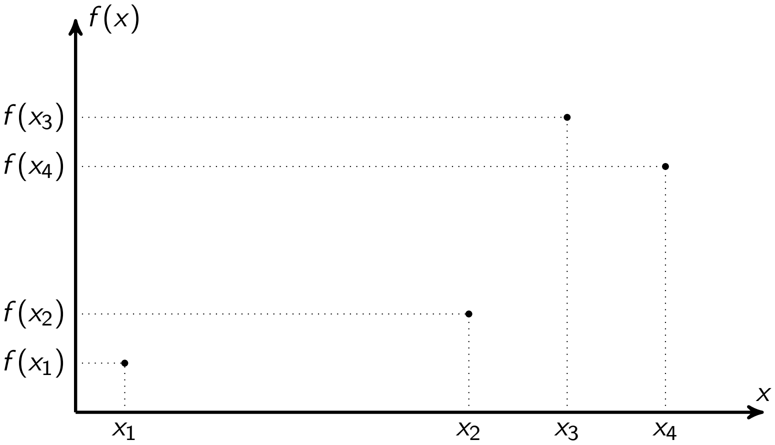
Parallel Gaussian Process Optimization with Upper Confidence Bound and Pure Exploration

Emile Contal
David Buffoni
Alexandre Robicquet
Nicolas Vayatis

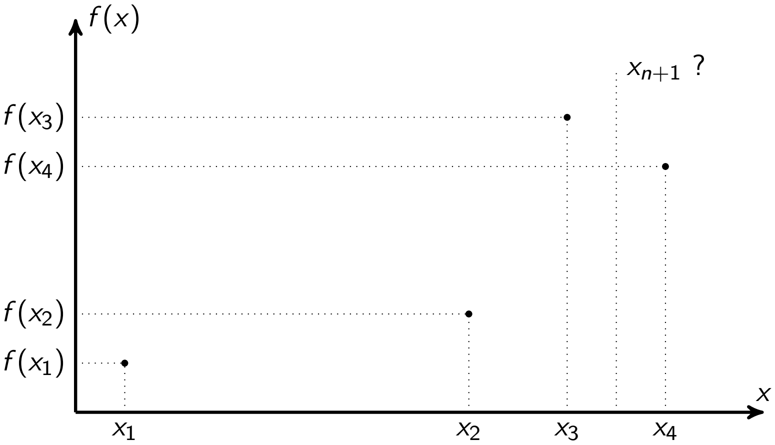
CMLA, ENS Cachan, France

September 25, 2013

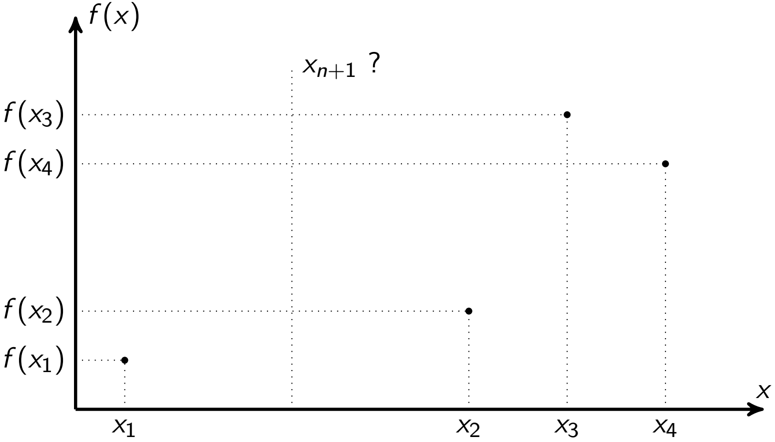
Motivating example: Sequential optimization



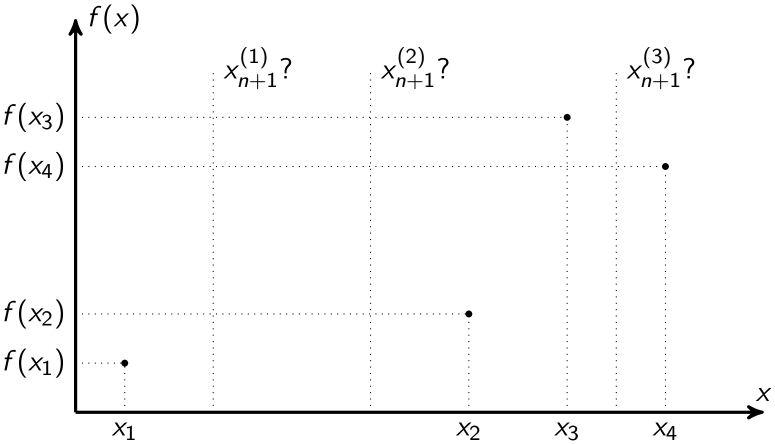
Motivating example: Sequential optimization



Motivating example: Sequential optimization



Motivating example: Batch optimization



Problem Statement

Setup

- Unknown $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \in \mathbb{R}^d$ compact and convex
- Find $x^\star = \operatorname{argmax}_{x \in \mathcal{X}} f(x)$
- At iteration t, query a batch of K locations $x_t^1, \dots, x_t^K \in \mathcal{X}$
- Observe the noisy evaluations of f, $y_t^1, \dots, y_t^K \in \mathbb{R}$,
  $y_t^k = f\left(x_t^k\right) + \epsilon_t^k$ where $\epsilon_t^k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$

Examples

- Heavy numerical experiment on a cluster with K cores
- Sensor placement with K sensors
- Laboratory experiment, ...

Bandit setting

Cumulative regret

- Exploration / Exploitation
- Batch cumulative regret: $R_T^K = \sum_{t=1}^{T} \Big(f(x^\star) - \max_{k \leq K} f(x_t^k) \Big)$
- Full cumulative regret: $R_{TK} = \sum_{t=1}^{T} \sum_{k=1}^{K} \Big(f(x^\star) - f(x_t^k) \Big)$

Gaussian Processes

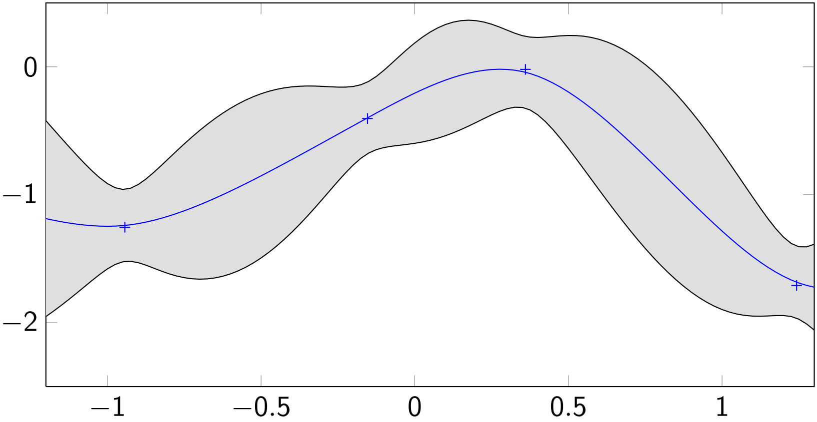
Definition

$f \sim \mathcal{GP}(m, k)$, with **mean** function $m : \mathcal{X} \to \mathbb{R}$ and **kernel** function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$, when for all $x_1, \dots, x_n$ the values $(f(x_1), \dots, f(x_n))$ are distributed as a **multivariate Gaussian** with mean and variance given by m and k.

Bayesian Inference (Rasmussen and Williams, 2005)

At iteration t, with observations $\mathbf{Y}_{X_t}$ at $X_t = \{x_1^1, \dots, x_t^K\}$, the posterior distribution $\Pr[f \mid \mathbf{Y}_{X_t}]$ is a Gaussian process with mean and covariance given by the Bayesian inference. At each point $x \in \mathcal{X}$, we can compute a **prediction** with $\mu_{t+1}(x)$ and **uncertainty** with $\sigma_{t+1}^2(x)$.

Example: Bayesian inference with 4 observations



Upper and Lower Confidence Bounds

Definition

Fix $0 < \delta < 1$,

$$f_t^+(x) = \mu_t(x) + \sqrt{\beta_t \sigma_t^2(x)}$$
$$f_t^-(x) = \mu_t(x) - \sqrt{\beta_t \sigma_t^2(x)}$$

with $\beta_t = \mathcal{O}(\log \frac{t}{\delta})$ defined in Srinivas et al. (2012)

Property

$$\forall x \in \mathcal{X}, \forall t \geq 1,$$

$$f(x) \in \left[f_t^-(x), f_t^+(x)\right] \textit{ holds with probability at least } 1 - \delta$$

Relevant Region

Definition

The Relevant Region $\mathfrak{R}_t$ and the extended $\mathfrak{R}_t^+$ are defined by,

$$y_t^\bullet = \max_{x \in \mathcal{X}} f_t^-(x)$$

$$\mathfrak{R}_t = \left\{ x \in \mathcal{X} \mid f_t^+(x) \geq y_t^\bullet \right\}$$

$$\mathfrak{R}_t^+ = \left\{ x \in \mathcal{X} \mid \mu_t(x) + 2\sqrt{\beta_{t+1}\sigma_t^2(x)} \geq y_t^\bullet \right\}$$

Property

$\text{argmax}_{x \in \mathcal{X}} \widehat{f}_{t+1}^+(x) \in \mathfrak{R}_t^+$ with high probability

UCB and Pure Exploration

UCB policy

$x_t^1 \leftarrow \text{argmax}_{x \in \mathfrak{R}_t^+} \widehat{f}_t^+(x)$

Pure Exploration policy

For $2 \geq k \geq K$, $x_t^k \leftarrow \text{argmax}_{x \in \mathfrak{R}_t^+} \sigma_t^{(k)}(x)$

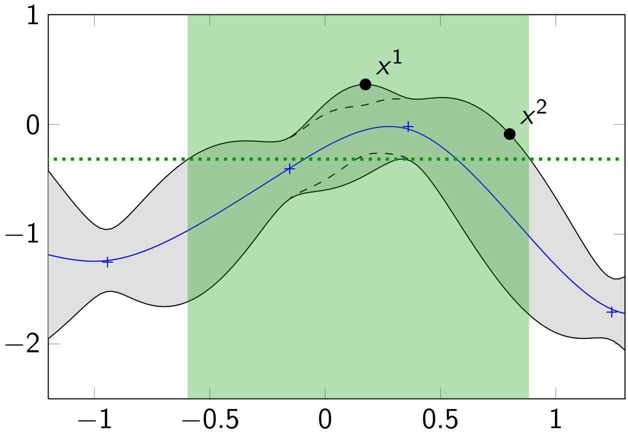
where $\sigma_t^{(k)}(x)$ is the updated deviation after having selected $x_t^1, \dots, x_t^{k-1}$

GP-UCB-PE

Algorithm 1: GP-UCB-PE

for $t = 1, 2, \dots$ **do**
- Compute μ_t and σ_t^2 with Bayesian inference on $y_1^1, \dots, y_{t-1}^K$
- Compute $\mathfrak{R}_t^+$
- $x_t^1 \leftarrow \mathrm{argmax}_{x \in \mathfrak{R}_t^+} \widehat{f}_t^+(x)$
- **for** $k = 2, \dots, K$ **do**
  - Update $\sigma_t^{(k)}$
  - $x_t^k \leftarrow \mathrm{argmax}_{x \in \mathfrak{R}_t^+} \sigma_t^{(k)}(x)$
- Query $\{x_t^k\}_{1 \le k \le K}$
- Observe $\{y_t^k\}_{1 \le k \le K}$

Example: GP-UCB-PE



Regret Bounds

Theorem

With $f \sim \mathcal{GP}(0,k)$ *and* $k(x,x) \leq 1$, *with high probability:*

$$R_T^K = \mathcal{O}\left(\sqrt{\gamma_{TK}\frac{T}{K}\log T}\right)$$

$$\textit{and } R_{TK} = \mathcal{O}\Big(\sqrt{\gamma_{TK}\,TK\log T}\Big)$$

Mutual Information

γ_{TK} is the maximum mutual information about f obtainable by a sequence of TK queries.

- For linear kernel, $\gamma_{TK} = \mathcal{O}(d\log TK)$
- For RBF kernel, $\gamma_{TK} = \mathcal{O}\big((\log TK)^{d+1}\big)$

Corollary

Corollary (Batch vs Sequential)

With $K \ll T$, the improvement of the parallel strategy over the sequential one is $\sqrt{K}$ with respect to R_T^K.

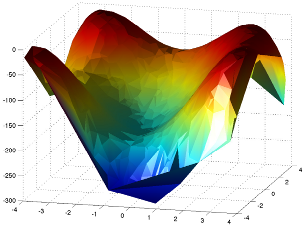
Remark (GP-BUCB, Desautels et al. (2012))

Compared to GP-BUCB, there is no need for an initialization phase. The improvement can be doubly exponential in the dimension d.

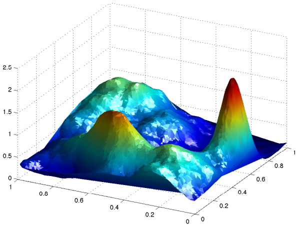
Experiments

Setup

- Competitors: GP-BUCB (Desautels et al. (2012)) and SM+UCB (Azimi et al. (2010))
- Assessment: 3 synthetic problems and 3 real applications

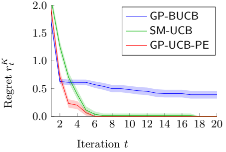


(a) Himmelblau

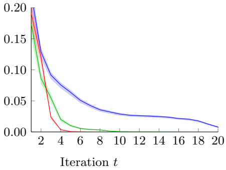


(b) Gaussian Mixture

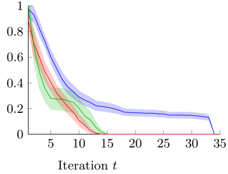
Results: mean instantaneous batch regret and confidence interval over 64 experiments



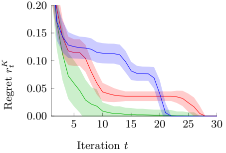
(a) Generated GP



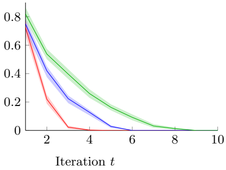
(b) Himmelblau



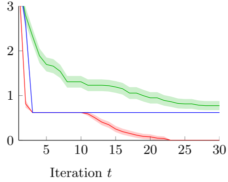
(c) Gaussian mixture



(d) Mackey-Glass



(e) Tsunamis



(f) Abalone

Conclusion

GP-UCB-PE

- Theoretical upper bounds on the cumulative regret
- Efficient in practice
- Easy to implement

Implementation

MATLAB source codes, documentation and data sets are available at `http://econtal.perso.math.cnrs.fr/software/`.

Azimi, J., Fern, A., and Fern, X. (2010). Batch bayesian optimization via simulation matching. In *Advances in Neural Information Processing Systems 24*, pages 109–117. Curran Associates, Inc.

Desautels, T., Krause, A., and Burdick, J. (2012). Parallelizing exploration-exploitation tradeoffs with gaussian process bandit optimization. In *Proceedings of the 29th International Conference on Machine Learning*, pages 1191–1198. icml.cc / Omnipress.

Rasmussen, C. E. and Williams, C. (2005). *Gaussian Processes for Machine Learning*. MIT Press.

Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. W. (2012). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5):3250–3265.