Bayesian global optimization
Project III (MATH3382)

Advisor: Georgios P. Karagiannis

Academic year 2017-2018 @ Durham University

Description

Bayesian global optimization (BGO) is a machine learning statistical methodology which aims at optimizing objective functions which are extremely expensive to be directly evaluated many times, and whose analytic expression and derivatives may not be available. By expensive, we mean in time, money, etc...

BGO is often used as a tool for the analysis of large-scale real world applications in climatology, engineering, chemistry, etc.... It can address problems such as inverse problems, computer model calibration, standard parameter estimation, etc....

BGO methods are applicable in scenarios where the objective function (to be optimized) is not available in closed-form, but it can be evaluated point-wisely at given input values. It is particularly useful when these evaluations are costly and possibly contaminated by noise.

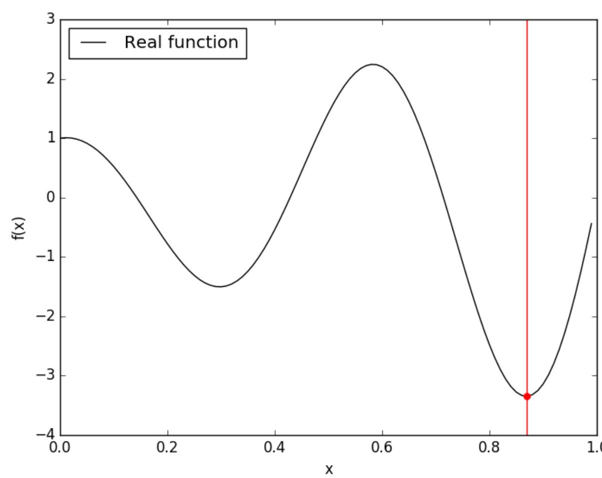
Central to BGO is the idea of building a probabilistic surrogate model for the objective function and using it to define a utility function called 'information acquisition function' (IAF) whose role is to guide the search for the optimum value. Given an IAF, BGO iterates between making the observation with the largest expected IAF and rebuilding the probabilistic surrogate until a convergence criterion is met.

- Consider a function $f(x) = \exp(1.4x)\cos(3.5\pi x)$ with minimum $f_{\min} \approx -3.3470$ at location $x_{\min} \approx 0.8686$.

  on the right $\longrightarrow$

- Pretend, we do not know the equation of $f(\cdot)$.
- Assume that we wish itsto find its minimum and minimum location by BGO.
- To see the BGO in action, discovering the global minimum and recovering the real function, click [HERE].

  Red dots: the samples; Blue lines: the recovered function and error bounds; red line: the acquisition function.



Objectives:

The objective of the project is to study several aspects of BGO. For instance, BGO variations to address problems involving:

- optimization under constraints
- time-dependent models
- heteroscedastisity
- non-stationarity
- multi-fidelity models

Pre-requisites

- Statistical Concepts II
- Knowledge of a programming language (required)

Co-requisites

- Bayesian Statistics III/IV (recommended, but not required)

References

- Brochu, E., Cora, V. M., & De Freitas, N. (2010). A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:1012.2599. [LINK]
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. In Advances in neural information processing systems (pp. 2951-2959) [LINK]
- Mockus, J. (1975). On Bayesian methods for seeking the extremum. In Optimization Techniques IFIP Technical Conference (pp. 400-404). Springer Berlin Heidelberg. [LINK]
- Williams, C. K., & Rasmussen, C. E. (2006). Gaussian processes for machine learning. the MIT Press, 2(3), 4. [LINK]

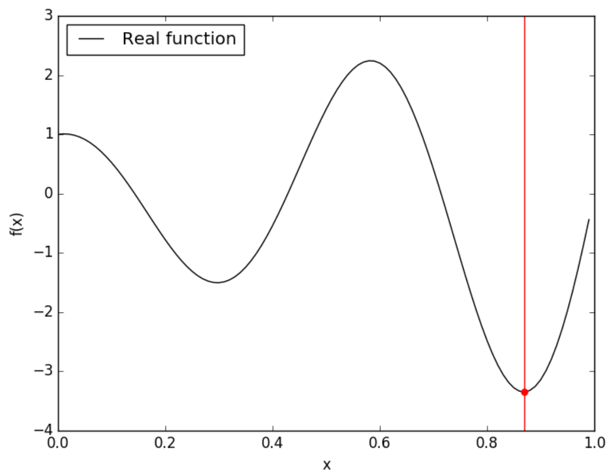
Example

Let $f(\cdot)$ be a function with equation

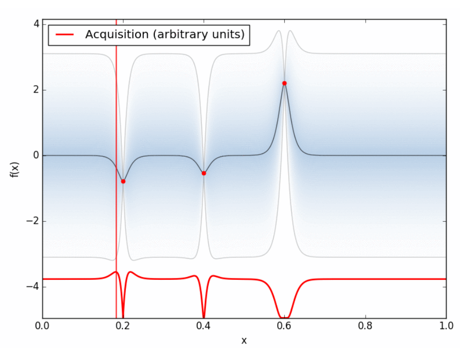
$$f(x) = \exp(1.4x)\cos(3.5\pi x)$$

with minimum $f_{\min} \approx -3.3470$ at location $x_{\min} \approx 0.8686$. Pretend the equation of $f(\cdot)$ and its minimum are unknown.

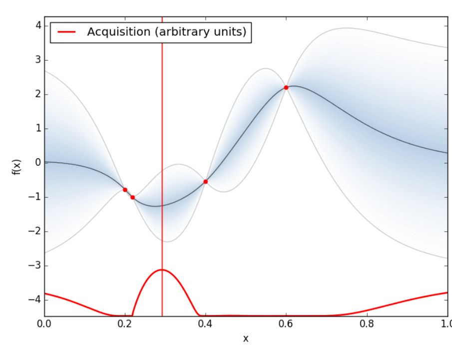
Assume we are interested in learning the minimum and the location of minimum of $f(\cdot)$.



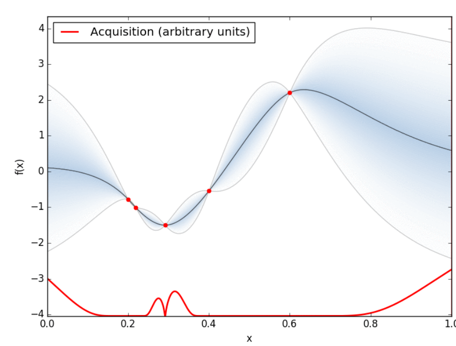
The BGO in action, discovering the global minimum and recovering the real function is shown below.
Red dots: the samples;
Blue lines: the recovered function and error bounds
Red line: the acquisition function



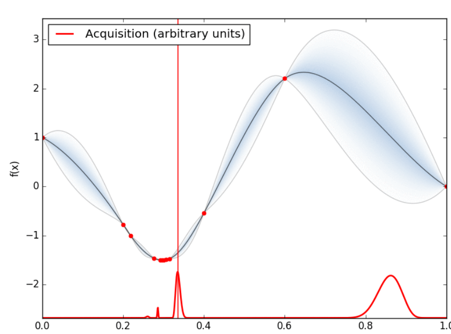
(a) iteration 1



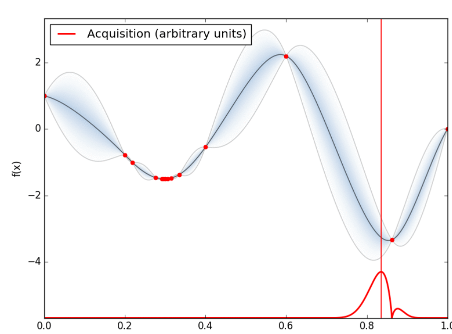
(b) iteration 2



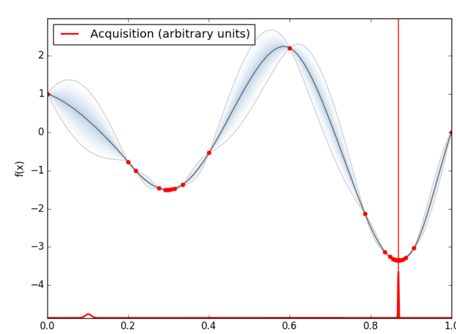
(c) iteration 3



(d) iteration 10



(e) iteration 12



(f) iteration 31