Conditional Quantile Optimization via Branch-and-Bound Strategies

Léonard Torossian
INRA (MIAT) - University of Toulouse (Institute of Mathematics)

Supervisor(s): Dr. Robert Faivre (INRA Toulouse), Prof. Aurélien Garivier (ENS-Lyon) and Dr. Victor Picheny (Prowler.io, Cambridge, UK)

Ph.D. expected duration: Nov. 2016 - Oct. 2019

Adress: INRA, 24 chemin de Borde Rouge, 31320 Auzeville-Tolosane

Email: leonard.torossian@inra.fr

Abstract:

We propose two branch-and-bound algorithms to optimize the conditional quantiles of *stochastic black boxes*. We consider systems that can be modeled as: $f : \mathcal{X} \times \Omega \to \mathbb{R}$, with f a reward or cost function, $\mathcal{X} \subset \mathbb{R}^D$ a design space and Ω a stochastic space. Contrary to *deterministic black boxes*, at a fixed x, the output is a random variable $Y_x = f(x, \omega)$ that follows an unknown distribution $\mathbb{P}(Y|X = x)$. We consider the classical setting of computer experiments: the function is only accessible through pointwise evaluations $f(x, \omega)$ and the gradient of any functional of f is unknown. Additionally, we assume that the variance of $\mathbb{P}(Y|X = x)$ may vary with respect to x (heteroscedasticity) and that $\mathbb{P}(Y|X = x)$ does not belong to any specific parametric class.

The conditional quantile function of order τ is defined as $q_\tau(x) = \inf\{q : F(q|X = x) \geq \tau\},\ \tau \in (0, 1)$, where $F(\cdot|X = x)$ is the cumulative distribution function of $\mathbb{P}(Y|X = x)$, and we denote its supremum (that is assumed to be a maximum) as $q_\tau(x^*)$. Given a finite evaluation budget n, the goal of the presented algorithms is to propose a value $x(n)$ such that the *simple regret*

$$r_n = q_\tau(x^*) - q_\tau(x(n))$$

is as small as possible. To do so, the algorithms use a sequential strategy $x_1, \ldots, x_n$ that balances between exploration and intensification. Once the budget is over, the value with the best theoretical guarantees $x(n)$ is returned.

Classical bandit-based approaches [2] rely on an upper confidence bound (UCB) of the objective function instead of a simple estimator, that is, a surrogate function larger than the objective supremum with high probability. Computing an UCB everywhere on a continuous input space may be difficult. To facilitate the computation and guide the sampling strategy, the algorithms that we consider are based on recursive partitionings of $\mathcal{X}$ represented by hierarchical partition trees $\mathcal{T}$. The association between UCB and hierarchical trees has been widely used to optimize the conditional expectation [2, 5]. However, to our knowledge, we are the first to rely on this type of strategies in order to optimize conditional quantiles over continuous spaces.

We propose here two algorithms: Quantile Optimistic Optimization (QOO), which is an adaptation of the Deterministic Optimistic Optimization (DOO) [4], and Quantile Hierarchical Optimization (QHO), which is inspired from the Hierarchical Optimistic Optimization (HOO) [2]. Following the work of [6] on discrete problems, both rely on a deviation inequality for the empirical quantile [3] to build an UCB. The UCB function is designed to favor the leaves most likely to contain x^* in order to create an accurate estimator near the possible optimal points, while also favoring the leaves that have been least explored.

The principle of QOO is as follow. Starting from an initial partitioning $\mathcal{T}_1$ of $\mathcal{X}$, at each step t ($1 \leq t \leq n$) QOO computes an UCB for all the leaves. Then the leaf with the highest UCB is

selected and f is evaluated for an x chosen inside that leaf. If the quantile estimate associated to the leaf is accurate enough then the leaf is expanded into K children. Determining if the estimator associated to the leaf is accurate enough is a key point of QOO: this is achieved by associated a lower confidence bound (rely again on the deviation inequality [3]) to the UCB. Finally, $x(n)$ is chosen in the deepest node among those that have been expanded.

QHO differs both in the UCB computation and the tree growing process. Contrary to QOO, at each time t, QHO finds an "optimistic" path from the root to the leaf. That process implies the computation of an UCB at each depth h, based on all the subsequent children. Once a leaf is reached, QHO expands it immediately and f is evaluated in the K new leaves. At the end, $x(n)$ is chosen uniformly at random among the points that have been evaluated.

DOO only samples at the center of each leaf while HOO samples randomly inside the leaf, which leads in practice to different observation sets (with our without repeated observations). In our framework, both QOO and QHO can work with either sampling strategy.

Using analysis tools from the bandit literature [5, 1], we prove non asymptotic guarantees for our algorithms, assuming only a smoothness relative to the optimal point:

$$q_\tau(x^*) - q_\tau(x) \leq A \left\| x^* - x \right\|^\beta, \quad \beta > 0, \quad A > 0.$$

We finally provide an empirical analysis that illustrates the respective merits of our algorithms. In particular, we discuss the benefits and drawbacks of the use of repetitions in the sampling strategies.

References

[1] Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.

[2] Sébastien Bubeck, Rémi Munos, Gilles Stoltz, and Csaba Szepesvári. X-armed bandits. *Journal of Machine Learning Research*, 12(May):1655–1695, 2011.

[3] Pascal Massart. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of Probability*, pages 1269–1283, 1990.

[4] Rémi Munos. Optimistic optimization of a deterministic function without the knowledge of its smoothness. In *Advances in neural information processing systems*, pages 783–791, 2011.

[5] Rémi Munos. From bandits to monte-carlo tree search: The optimistic principle applied to optimization and planning. *Foundations and Trends® in Machine Learning*, 7(1):1–129, 2014.

[6] Balazs Szorenyi, Róbert Busa-Fekete, Paul Weng, and Eyke Hüllermeier. Qualitative multi-armed bandits: A quantile-based approach. In *32nd International Conference on Machine Learning*, pages 1660–1668, 2015.

Short biography – I graduated from Université Pierre et Marie Curie in 2016 with a MSc in Modelling and Optimization. I started my PhD at INRA and IMT in November 2016 co-directed by R. Faivre, V. Picheny and A. Garivier. My work is funded by MIA and the Occitanie region. My PhD subject is about metamodeling and robust optimization of stochastic black boxes, I aim to build a tool able to take optimal decisions under risk aversion.