Bayesian multi-objective optimization with noisy evaluations
Bruno Barracosa, Julien Bect, Héloïse Dutrieux Baraffe, Juliette Morin,
Josselin Fournel, Emmanuel Vazquez

To cite this version:
Bruno Barracosa, Julien Bect, Héloïse Dutrieux Baraffe, Juliette Morin, Josselin Fournel, et al.. Bayesian multi-objective optimization with noisy evaluations. MASCOT PhD student 2020 Meeting, Sep 2020, Grenoble, France. hal-03022267

HAL Id: hal-03022267
https://hal-centralesupelec.archives-ouvertes.fr/hal-03022267
Submitted on 24 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
BAYESIAN MULTI-OBJECTIVE OPTIMIZATION WITH NOISY EVALUATIONS

1. CONTEXT

- Multi-objective optimization of the parameters of a planning strategy for the multi-year planning of the electricity distribution grid [DUT15]
- A stochastic black box provides noisy evaluation results of the objective functions $f_1, ..., f_d$ defined on a discrete search domain $X \subset \mathbb{R}^d$
- Previous $n$ evaluations at $X = \{X_1, ..., X_n\}$ assumed:
  - $f_i = f_i(X_i) + \varepsilon_i$, where $\varepsilon_i$ is zero-mean random variable

2. OBJECTIVE

Estimate the Pareto-optimal solutions (or Pareto set $\Gamma$) of the problem:

$$x^* = \text{argmin}_{x \in \mathbb{R}^d} f(x)$$

Defined as:

$$\Gamma = \{x \in \mathbb{R}^d : \exists x' \in \mathbb{R}^d \text{ such that } f(x') < f(x)\}$$

where $<$ stands for the Pareto domination rule:

$$y = (y_1, ..., y_n) \prec y' = (y'_1, ..., y'_m) \iff \forall i \leq n, y_i \leq y'_i \land \exists j \leq m, y_j < y'_j$$

Illustration of the Pareto dominance rule

- $y, y_2, y_3$ are non-dominated points
- $y_1$ is only dominated by $y_3$
- $y_4$ is dominated by all other points

3. BAYESIAN OPTIMIZATION

- Define a probabilistic model for each $f$ conditional on previous observations
- Use a sampling criterion to select new evaluation points

Provides estimate of $f$ and a measure of uncertainty of the estimation

**PARETO-OPTIMAL ESTIMATES**

Built from the estimates of $f$

PROPOSED APPROACH

- Replace the multi-objective problem by the minimization of a single augmented Tchebycheff function [KNO06]:

$$f(x) = \max_i \left[ \omega_i f_i(x) \right] + \rho \sum_i \omega_i f_i(x)$$

- At each iteration, generate random weights $\omega_i$ and apply this function to the $n$ previous observations:

$$\hat{Z} = f(x)$$

- Assume a homoscedastic Gaussian noise model and fit to $Z = (\tilde{Z}_1, ..., \tilde{Z}_n)$ a Gaussian Process model $\xi_i$ with parameters estimated by maximum likelihood

- We use the Knowledge Gradient (KG) criterion [FRA09] to select new point $X_n+1$ based on previous observations. The idea is to identify a point that is expected to reduce the minimum of the posterior mean of $\xi$:

$$X_{n+1} = \text{argmax}_{x \in \mathbb{R}^d} \text{KG}(x)$$

with:

$$\text{KG}(x) = \min_{x' \in \mathbb{R}^d} \mathbb{E}[f(x') | z] - \mathbb{E}[\min_{x' \in \mathbb{R}^d} f(x') | z_{n+1}]$$

where $z_{n+1}$ denotes a new observation of $\xi$ at $x$

- Update the model with new observation and iterate until stopping criterion is met

4. NUMERICAL EXPERIMENTS

- Compare the proposed approach (PA) with random selection of points in a bi-dimensional bi-objective problem
- Compare use of batches of 1, 4 or 10 evaluations

Hypervolume difference between the Pareto front and estimates

Hypervolume distance between the Pareto front and estimates

Average performance results over 100 simulations for four optimization approaches: random approach (red) and the proposed approach with batches of 1 (black), 4 (blue) or 10 (magenta) evaluations

5. OPEN QUESTIONS

- ‘Ideal’ batch of evaluations?
- Performance comparison of the proposed approach to other methods in the literature?

REFERENCES


Contacts:
Bruno BARACOSA (EDF R&D, L2S, CentraleSupélec)
Julien BECT (L2S, CentraleSupélec)
Hélène DUTRIEUX BARAFFE (EDF R&D)
Juliette MORIN (EDF R&D)
Joséfin FOURNEL (EDF R&D)
Emmanuel VAZQUEZ (L2S, CentraleSupélec)
*bruno.b_goals@edf.fr