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 \( f_1, \ldots, f_d \) defined on a discrete search domain \( X \subset \mathbb{R}^d \)

• Previous \( n \) evaluations at \( X = (X_1, \ldots, X_n) \) assumed:

\[
\begin{align*}
X_i \in X & \quad \text{Stochastic black box} \\
Z_1, \ldots, Z_q & \quad \text{Previous evaluations at} \ X
\end{align*}
\]

where \( \xi_i \) are zero-mean random variables

2. OBJECTIVE

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

\[
x^* = \arg\min_{x \in \mathbb{X}} f(x) \quad \text{subject to} \quad g(x) \leq 0, \ h(x) = 0
\]

Defined as:

\[
\Gamma = \{ x \in \mathbb{X} : \exists x' \in \mathbb{X} \text{ such that } f(x') < f(x) \}
\]

where \( \prec \) stands for the Pareto domination rule:

\[
y_1 \prec y_2 \quad \text{iff} \quad \forall i \leq q, y_i \leq y_i' \quad \text{and} \quad \exists j \leq q, y_j < y_j'
\]

Illustration of the Pareto domination rule

• \( y_1, y_2, y_3 \) are non-dominated points

• \( y_2' \) is only dominated by \( y_2 \)

• \( y_3' \) 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]:

\[
f(x) = \max \left[ \omega_1 f_1(x) \right] + \rho \sum_{j=1}^d \omega_j f_j(x), \quad \sum_1^n \omega_j = 1, \rho > 0
\]

• At each iteration, generate random weights \( \omega_j \) and apply this function to the \( n \) previous observations:

\[
Z_i = \max \left[ \omega_1 Z_{ij} \right] + \rho \sum_{j=1}^d \omega_j Z_{ij}, \quad i = 1, \ldots, n
\]

• Assume a homoscedastic Gaussian noise model and fit to \( Z = (Z_1, \ldots, Z_n) \) 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 \( f \): \( X_{n+1} = \arg\max_{x \in \mathbb{X}} KG (x) \)

with:

\[
KG(x) = \min_{x' \in \mathbb{X}} E [ f(x') ] - E [ \min_{x' \in \mathbb{X}} f(x') | Z_{1:n} ]
\]

where \( Z_{1:n} \) denotes a new observation of \( \xi_i \) at \( x \)

4. NUMERICAL EXPERIMENTS

• Compare use of batches of 1, 4 or 10 evaluations

5. OPEN QUESTIONS

• 'Ideal' batch of evaluations?

• Performance comparison of the proposed approach to other methods in the literature?

REFERENCES


