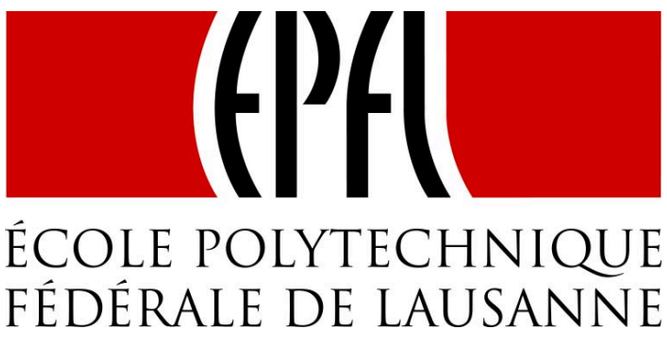
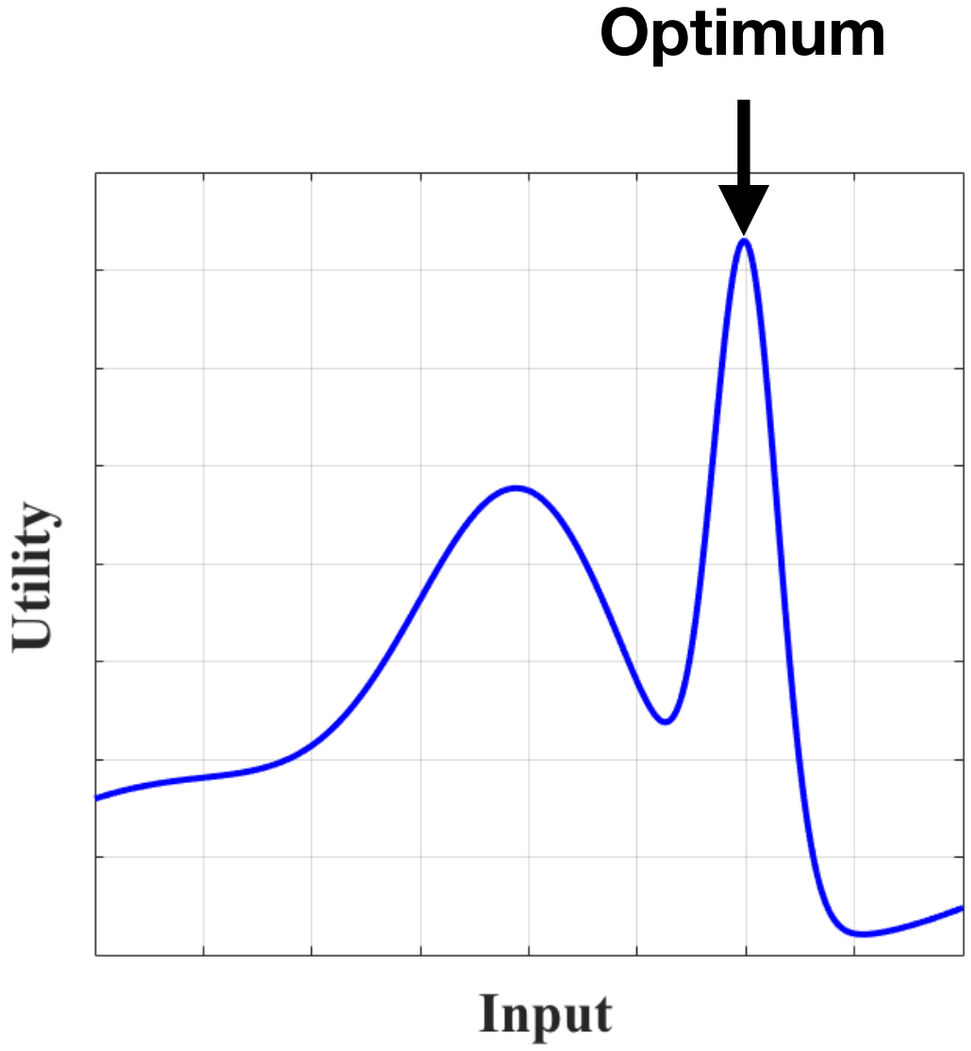
Adversarially Robust Optimization with Gaussian Processes

Ilija Bogunovic, Jonathan Scarlett, Stefanie Jegelka, Volkan Cevher

Conference on Neural Information Processing Systems (Dec 2018)



Gaussian Process Optimization



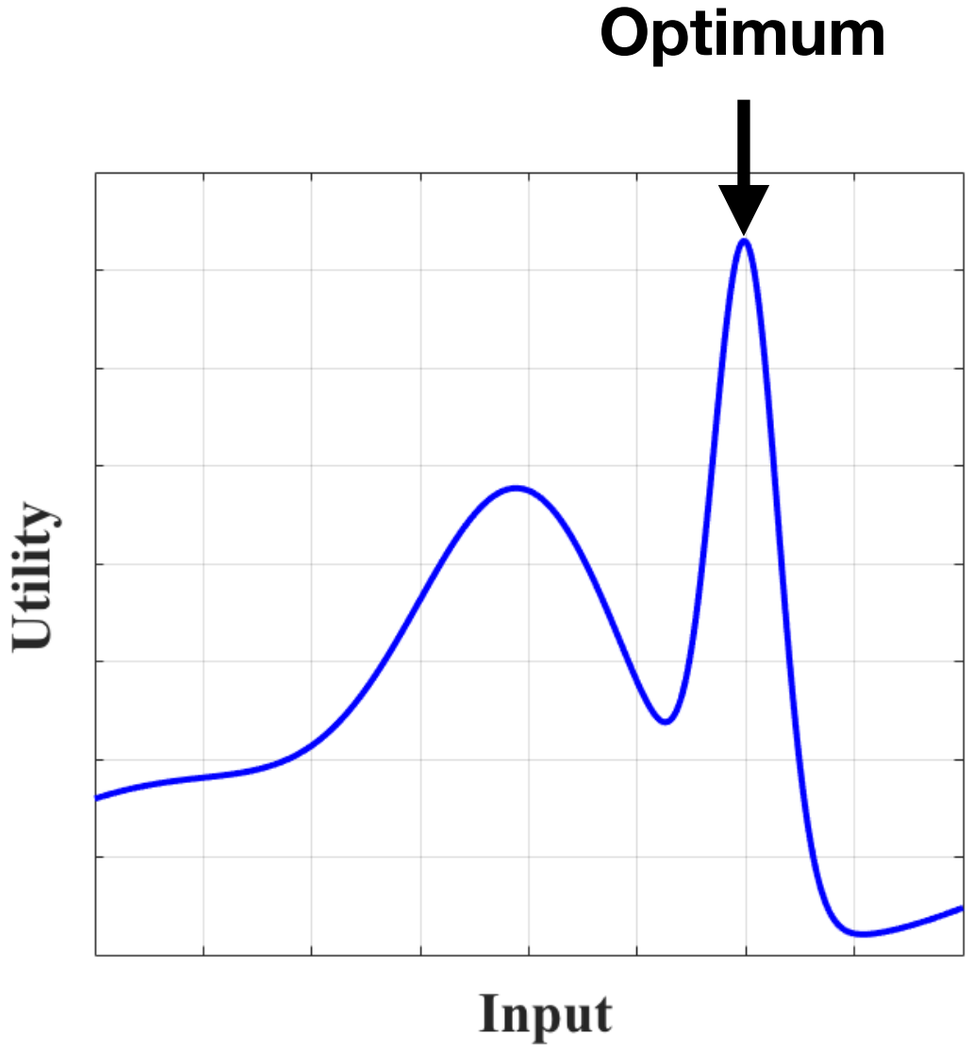
Non-robust problem:

$$\boldsymbol{x}^* = \arg\max_{\boldsymbol{x} \in D \subset \mathbb{R}^d} f(\boldsymbol{x})$$

Setting: GP/Bayesian optimization

- Unknown utility function f, modeled by Gaussian Process $f \sim \mathrm{GP}(\mu, \kappa)$
- Sequentially query the unknown function f
- Noisy and expensive point evaluations

Adversarially Robust GP Optimization



Robust problem:

$$x^* = \arg\max_{x \in D \subset \mathbb{R}} \min_{\delta \in \Delta_\epsilon(x)} f(x+\delta)$$

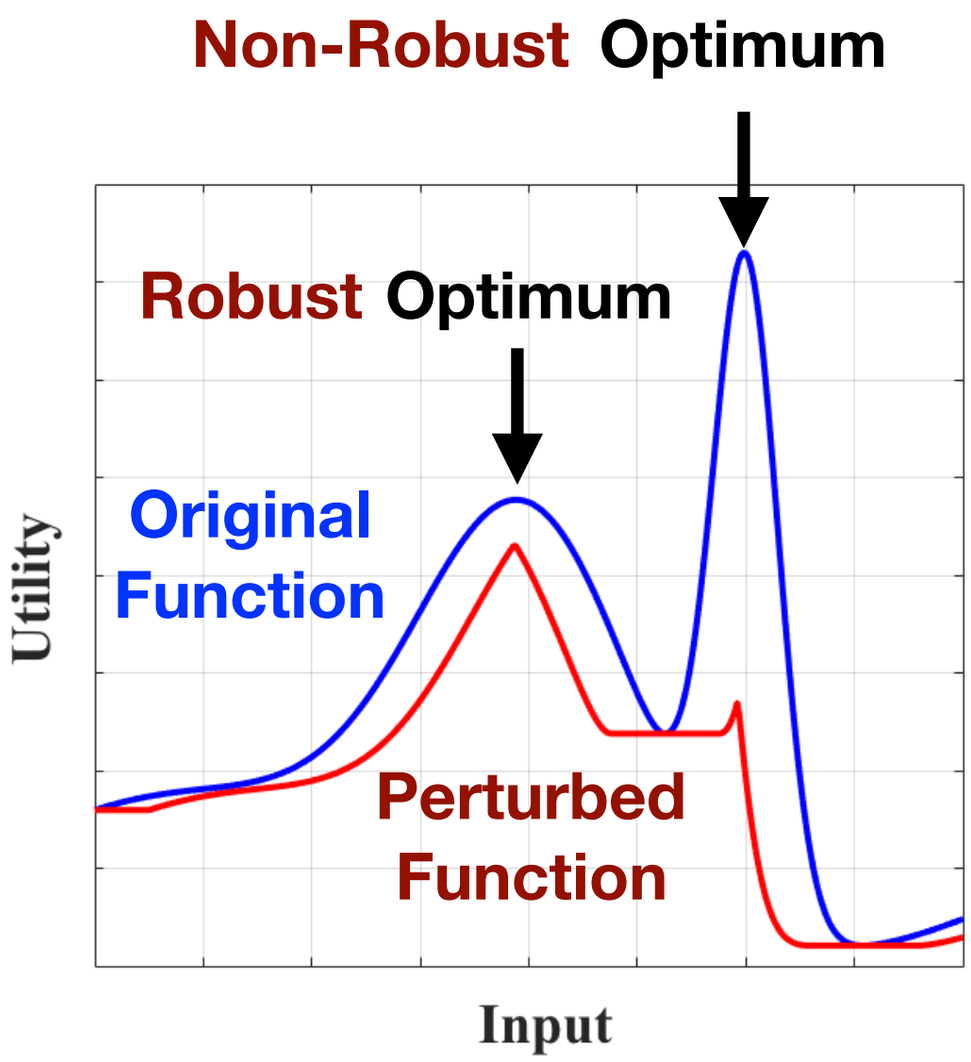
Set of input perturbations:

$$\Delta_\epsilon(x) = \{x' - x \ : \ \mathrm{dist}(x, x') \leq \epsilon\}$$

Setting:

- Unknown utility function f, modeled by Gaussian Process $f \sim \mathrm{GP}(\mu, \kappa)$
- Sequentially query the unknown function f
- Noisy and expensive point evaluations

Adversarially Robust GP Optimization



Robust problem:

$$x^* = \arg\max_{x \in D \subset \mathbb{R}} \min_{\delta \in \Delta_\epsilon(x)} f(x+\delta)$$

Set of input perturbations:

$$\Delta_\epsilon(x) = \{x' - x \ : \ \mathrm{dist}(x, x') \leq \epsilon\}$$

Motivation: adversarial attack, implementation errors, etc.

Setting:

- Unknown utility function f, modeled by Gaussian Process $f \sim \mathrm{GP}(\mu, \kappa)$
- Sequentially query the unknown function f
- Noisy and expensive point evaluations

Robust Algorithm: StableOpt

Non-robust BO methods:

Thompson [Thompson ’33]

PI [Kushner’64]

EI [Mockus *et al.’78*]

GP-UCB [Srinivas *et al.’11*]

ES [Henning *et al.*’12]

GP-UCB-PE [Contal *et al.’13*]

BamSOO [Wang *et al.*’14]

PES [Hernandez-Lobato *et al.*’14]

MRS [Metzen’16]

GLASSES [Gonzalez *et al.*’15]

OPES [Hoffman & Ghahramani’15]

TruVaR [Bogunovic *et al.*'16]

MES [Wang & Jegelka’17]

FITBO [Ru *et al.*’18]

KG [Wu *et al.*’17]

the list goes on…

Robust Algorithm: StableOpt

Robust algorithm: **StableOpt**

Round t :

- Choose: $\tilde{\boldsymbol{x}}_t = \underset{\boldsymbol{x} \in D}{\operatorname{argmax}} \underset{\boldsymbol{\delta} \in \Delta_\epsilon(\boldsymbol{x})}{\min} \operatorname{ucb}_{t-1}(\boldsymbol{x} + \boldsymbol{\delta})$

Non-robust BO methods:

Thompson [Thompson '33]

PI [Kushner'64]

EI [Mockus *et al.'78*]

GP-UCB [Srinivas *et al.'11*]

ES [Henning *et al.*'12]

GP-UCB-PE [Contal *et al.'13*]

BamSOO [Wang *et al.*'14]

PES [Hernandez-Lobato *et al.*'14]

MRS [Metzen'16]

GLASSES [Gonzalez *et al.*'15]

OPES [Hoffman & Ghahramani'15]

TruVaR [Bogunovic *et al.*'16]

MES [Wang & Jegelka'17]

FITBO [Ru *et al.*'18]

KG [Wu *et al.*'17]

the list goes on…

Robust Algorithm: StableOpt

Robust algorithm: **StableOpt**

Round t :

- Choose: $\tilde{\boldsymbol{x}}_t = \underset{\boldsymbol{x}\in D}{\operatorname{argmax}} \underset{\boldsymbol{\delta}\in\Delta_\epsilon(\boldsymbol{x})}{\min} \operatorname{ucb}_{t-1}(\boldsymbol{x}+\boldsymbol{\delta})$
- Select: $\boldsymbol{\delta}_t = \underset{\boldsymbol{\delta}\in\Delta_\epsilon(\tilde{\boldsymbol{x}}_t)}{\operatorname{argmin}} \operatorname{lcb}_{t-1}(\tilde{\boldsymbol{x}}_t+\boldsymbol{\delta})$

Non-robust BO methods:

Thompson [Thompson ’33]

PI [Kushner’64]

EI [Mockus *et al.’78*]

GP-UCB [Srinivas *et al.’11*]

ES [Henning *et al.*’12]

GP-UCB-PE [Contal *et al.’13*]

BamSOO [Wang *et al.*’14]

PES [Hernandez-Lobato *et al.*’14]

MRS [Metzen’16]

GLASSES [Gonzalez *et al.*’15]

OPES [Hoffman & Ghahramani’15]

TruVaR [Bogunovic *et al.*'16]

MES [Wang & Jegelka’17]

FITBO [Ru *et al.*’18]

KG [Wu *et al.*’17]

the list goes on…

Robust Algorithm: StableOpt

Robust algorithm: **StableOpt**

Round t :

- Choose: $\tilde{\boldsymbol{x}}_t = \underset{\boldsymbol{x}\in D}{\operatorname{argmax}} \ \underset{\boldsymbol{\delta}\in\Delta_\epsilon(\boldsymbol{x})}{\min} \ \mathrm{ucb}_{t-1}(\boldsymbol{x}+\boldsymbol{\delta})$
- Select: $\boldsymbol{\delta}_t = \underset{\boldsymbol{\delta}\in\Delta_\epsilon(\tilde{\boldsymbol{x}}_t)}{\operatorname{argmin}} \ \mathrm{lcb}_{t-1}(\tilde{\boldsymbol{x}}_t+\boldsymbol{\delta})$
- Observe noisy function value at $\tilde{\boldsymbol{x}}_t+\boldsymbol{\delta}_t$

Non-robust BO methods:

Thompson [Thompson '33]

PI [Kushner'64]

EI [Mockus *et al.'78*]

GP-UCB [Srinivas *et al.'11*]

ES [Henning *et al.*'12]

GP-UCB-PE [Contal *et al.'13*]

BamSOO [Wang *et al.*'14]

PES [Hernandez-Lobato *et al.*'14]

MRS [Metzen'16]

GLASSES [Gonzalez *et al.*'15]

OPES [Hoffman & Ghahramani'15]

TruVaR [Bogunovic *et al.*'16]

MES [Wang & Jegelka'17]

FITBO [Ru *et al.*'18]

KG [Wu *et al.*'17]

the list goes on…

Theoretical Result

Theorem:

StableOpt guarantees that if

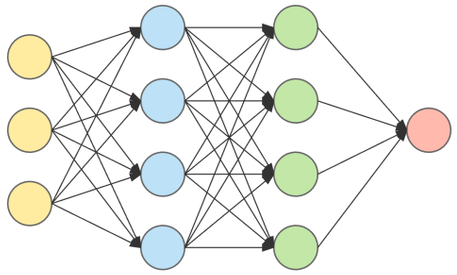
$$T \gtrsim \frac{\gamma_T}{\eta^2}$$

then the reported point $\boldsymbol{x}^{(T)}$ satisfies the following w.h.p.:

$$\min_{\boldsymbol{\delta} \in \Delta_\epsilon(\boldsymbol{x}^{(T)})} f(\boldsymbol{x}^{(T)} + \boldsymbol{\delta}) \geq \max_{\boldsymbol{x} \in D \subset \mathbb{R}} \min_{\boldsymbol{\delta} \in \Delta_\epsilon(\boldsymbol{x})} f(\boldsymbol{x} + \boldsymbol{\delta}) \; - \; \eta,$$

where
- T : Total number of points queried
- η : Target accuracy
- γ_T : Kernel-dependent information quantity

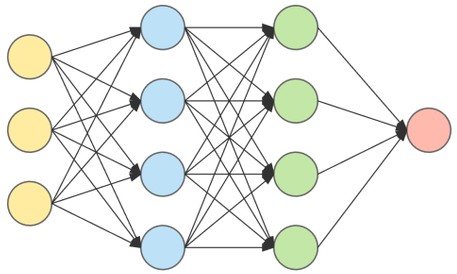
Variations



Robustness to unknown parameters:

- Goal: Choose $\boldsymbol{x}$ robust to different $\boldsymbol{\theta}$, $\max_{\boldsymbol{x}\in D} \min_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{x},\boldsymbol{\theta})$
- Application: Tuning hyperparameters robust to different data types

Variations



Robustness to unknown parameters:

- Goal: Choose $\boldsymbol{x}$ robust to different $\boldsymbol{\theta}$, $\max_{\boldsymbol{x}\in D} \min_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{x},\boldsymbol{\theta})$
- Application: Tuning hyperparameters robust to different data types

Robust group identification: Input space is partitioned into groups

G_1 G_2 G_k

- Goal: Identify the group with the highest worst-case function value

$$\max_{G\in\mathcal{G}} \min_{\boldsymbol{x}\in G} f(\boldsymbol{x})$$

- Application: Robust group movie recommendation

Adversarially Robust Optimization with Gaussian Processes

Ilija Bogunovic, Jonathan Scarlett, Stefanie Jegelka, Volkan Cevher

Poster #24

Wed Dec 5th 05:00 -- 07:00 PM @ Room 210 & 230 AB

European Research Council
Established by the European Commission

Adversarially Robust Optimization with Gaussian Processes

Ilija Bogunovic
LIONS, EPFL
ilija.bogunovic@epfl.ch

Jonathan Scarlett
National University of Singapore
scarlett@comp.nus.edu.sg

Stefanie Jegelka
MIT CSAIL
stefje@mit.edu

Volkan Cevher
LIONS, EPFL
volkan.cevher@epfl.ch

Abstract

In this paper, we consider the problem of Gaussian process (GP) optimization with an added robustness requirement: The returned point may be perturbed by an adversary, and we require the function value to remain as high as possible even after this perturbation. This problem is motivated by settings in which the underlying functions during optimization and implementation stages are different, or when one is interested in finding an entire region of good inputs rather than only a single point. We show that standard GP optimization algorithms do not exhibit the desired robustness properties, and provide a novel confidence-bound based algorithm StableOpt for this purpose. We rigorously establish the required number of samples for StableOpt to find a near-optimal point, and we complement this guarantee with an algorithm-independent lower bound. We experimentally demonstrate several potential applications of interest using real-world data sets, and we show that StableOpt consistently succeeds in finding a stable maximizer where several baseline methods fail.

1 Introduction

Gaussian processes (GP) provide a powerful means for sequentially optimizing a black-box function f that is costly to evaluate and for which noisy point evaluations are available. Since its introduction, this approach has successfully been applied to numerous applications, including robotics [21], hyperparameter tuning [30], recommender systems [34], environmental monitoring [31], and more.

In many such applications, one is faced with various forms of uncertainty that are not accounted for by standard algorithms. In robotics, the optimization is often performed via simulations, creating a mismatch between the assumed function and the true one; in hyperparameter tuning, the function is typically similarly mismatched due to limited training data; in recommendation systems and several other applications, the underlying function is inherently time-varying, so the returned solution may become increasingly stale over time; the list goes on.

In this paper, we address these considerations by studying the GP optimization problem with an additional requirement of *adversarial robustness*: The returned point may be perturbed by an adversary, and we require the function value to remain as high as possible even after this perturbation. This problem is of interest not only for attaining improved robustness to uncertainty, but also for settings where one seeks a region of good points rather than a single point, and for other related max-min optimization settings (see Section 4 for further discussion).

32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada.