MIE498 – Undergraduate Research Thesis Proposal Bayesian Optimization using Information Gain Aqcuisition Functions

Ruifan Wu
Department of Mechanical & Industrial Engineering
University of Toronto
ruifan.wu@mail.utoronto.ca
Student ID: 1006702926

1 Introduction

Bayesian optimization (BO) is a sequential optimization methodology for locating the global minima of a general nonlinear function that may be corrupted by noise. The standard BO formulation incorporates the use of Gaussian Process (GP) models that are used to model the objective function. GP models offer the benefit of including error statistics which can be interpreted to represent how certain the model is about it's prediction. Additionally, BO incorporates the use of acquisition functions (search functions, infill criteria or policies) to determine a promising candidate point which can be used to evaluate the true objective function. In this thesis project, we aim to investigate the performance of information gain matrics – such as entropy search, predictive entropy search and max value entropy search – when used as the specified acquisition function for BO.

2 Objectives

The objectives of this thesis project are

- Construct entropy-based acquisition functions (i.e., Entropy Search, Predictive Entropy Search and Max Value Entropy Search) that can be used in a Bayesian optimization setting where we use the statistics from a GP posterior distribution to compute the search policy for ES, PES and/or MES.
- Compare the performance of these acquisition functions in a Bayesian Optimization setting against other common acquisition functions such as the probability of improvement, expected improvement, lower confidence bound and/or Thompson sampling.
- Extend this work to the setting of optimization under uncertainty if time permits where we have Gaussian process models for robustness metrics (optional).

3 Description of Activities

To achieve these objectives, I will be focusing on:

- Complete Literature Review on
  - Gaussian Process Modelling
  - Acquisition Functions in Bayesian Optimization Modeling
- Curate multiple datasets in different dimensions in Bayesian Optimization settings
- Developing implementations for various acquisition functions including

  - Entropy-based acquisition functions
    - Entropy Search
    - Max-value Entropy Search
    - Predictive Entropy Search
  - Common acquisition functions
    - Probability of Improvement
    - Expected Improvement
    - Lower Confidence Bound
    - Thompson Sampling
- Testing and evaluating different proposed methodologies in the curated datasets and explore ideas on improving entropy-based search's performance and efficiency in normal and under uncertainty scenarios.

4 Evaluation Breakdown and Percentage

- Review of the literature: 10%
- Development and description of theory: 20%
- Implementation: 20%
- Evaluation: 25%
- Summary of Finding: 25%

5 Work Schedule Breakdown

January:

- Onboarding with GPytorch and Pytorch by finishing GPyTorch Regression Tutorial
- Implement the $2D$ and nD Gaussian process models in GPytorch environment.

February:

- Implement and test information gain metrics that can be used together with GP models
- Design of experiments to compute the performance of the information gain metrics

March:

- Finish any implementation.
- Carry out experiments

April:

- Finalize thesis write-up and submit for department review.

Table 1: Work Schedule

Deliverable/ Task	Date/ Time Needed
Meeting	Weekly / 1-2 hrs
Reading related materials	Weekly / 6 hrs
Coding	Weekly / 5 hrs
Testing and evaluations	Weekly / 2 hrs
Final Thesis Report	A week before the last day of lectures / 40 hrs

For the semester: around 230 hours.