Stanford University

AA222/CS361: Engineering Design Optimization

Spring 2023

Prof. Mykel J. Kochenderfer • NVIDIA Auditorium • email: *mykel@stanford.edu*

PROJECT 3: GAUSSIAN PROCESSES **Due date: May 26, 2023 (5:00pm PT)**

In this project, you will perform optimization under uncertainty using a Gaussian Process (GP), a probabilistic surrogate model that represents a distribution over functions. Surrogate models predict the true objective function when the latter is expensive or difficult to evaluate. Unlike other surrogate models, Gaussian Processes also quantify the uncertainty in their predictions. The goal of this project is to build a GP surrogate model for the provided function evaluations and to use different exploration techniques to choose future design points.

1 Instructions

1.1 Setup

You may use any programming language and package you wish (e.g., libraries to fit Gaussian Processes and use them to make predictions). Some popular Gaussian Process libraries include `GaussianProcesses` in Julia and `sklearn.gaussian_process` in Python. Make sure you thoroughly refer to the documentation of each package and go through the appropriate processes to install the packages into your programming environment before you begin coding.

1.2 Objective

We will use a GP to minimize an objective function

$$f(x) = \frac{x^2 + 5\sin(2x)}{2}$$

after having made five initial observations at $x_{\text{obs}} \in \{-4, -2, 0, 2.5, 4\}$.

1.3 Part 1: Gaussian Process Fitting

In this part, we fit a Gaussian Process to noiseless observations of our objective function and investigate the effect of using different characteristic length scales on resulting predictions.

1. Evaluate the objective function at the provided observed design values.
2. Fit and visualize a Gaussian Process to the data using a squared exponential kernel

$$k(x, x') = \exp(-\frac{(x - x')^2}{2\ell^2}).$$

Repeat the following for characteristic length-scales $\ell = 0.5, 1, \text{and } 2$:

(a) Fit a Gaussian Process with the appropriate kernel to the five design points. Note that you can use any Gaussian Process package. You should take care that the package fixes the kernel hyperparameters and doesn't try to fit them.

(b) Extract the predicted mean and the 95% confidence interval (1.96 standard deviations) at prediction points $x_{\text{pred}} \in -4 : 0.01 : 4$.

(c) Plot the true function, the five function evaluations, the predicted mean function, and the confidence interval on one plot. Refer to Figure 15.5 in the textbook for an example.

Set axis limits to $x_{\text{lim}} \in [-4, 4]$ and $y_{\text{lim}} \in [-3, 13]$.

1.4 Part 2: Noisy Observations

In this part, we assume that function evaluations are actually noisy, having additive Gaussian noise drawn from $\mathcal{N}(0, \sigma = 0.2)$. The noisy observations evaluate to:

$$x_{\text{obs}}, y_{\text{obs}} \in \{(-4, 5.64), (-2, 3.89), (0, 0.17), (2.5, 0.92), (4, 10.49)\}$$

Fix the squared exponential kernel length scale to $\ell = 0.9$ and repeat the process from Part 1 to generate a noisy GP prediction plot. Refer to Figure 15.7 for an illustrative example.

Set axis limits to $x_{\text{lim}} \in [-4, 4]$ and $y_{\text{lim}} \in [-3, 13]$.

1.5 Part 3: Exploration Strategies

Finally, we use different exploration strategies to guide the selection of the next design point. In this part, we want to investigate the difference between choosing the next design point using:

- Prediction-based exploration (Section 16.1)
- Error-based exploration (Section 16.2)
- Lower confidence bound exploration (Section 16.3) on the 95% confidence interval ($\alpha = 1.96$)

Building on the fitted noisy GP in Part 2, for **each** of these three exploration strategies:

1. Determine and print the next (sixth) design point x according to the exploration strategy among the possible points $x_{\text{pred}} \in -4 : 0.01 : 4$.
2. Evaluate the objective function f at the new design point. Assume no noise.
3. Refit the GP including the new sixth point.
4. Plot the six design points, true function, and new predicted GP mean and 95% confidence interval. Set axis limits to $x_{\text{lim}} \in [-4, 4]$ and $y_{\text{lim}} \in [-3, 13]$. You can refer to the figures in Sections 16.1–3 as illustrative examples.
5. **For the LCB exploration strategy only** determine and print the following (seventh) design point

Finally, **write a one-paragraph comparison** of the different strategies.

2 Submission

Please submit a PDF document containing the following information:

- Part 1: Noiseless Gaussian Process Fitting (Refer to Fig. 15.5 in textbook)
  - Plot for $\ell = 0.5$.
  - Plot for $\ell = 1.0$.
  - Plot for $\ell = 2.0$.
- Part 2: Noisy Gaussian Fitting (Refer to Fig. 15.7 in textbook)
  - Plot for $\ell = 0.9$ with additive Gaussian noise.
- Part 3: Exploration strategies
  - Tabulated values of the separate function evaluations when using the different exploration strategies (2 evaluations for LCB).
  - Updated noisy Gaussian Process prediction plot after including the next design point from prediction-based exploration.

  - Updated noisy Gaussian Process prediction plot after including the next design point from error-based exploration.
  - Updated noisy Gaussian Process prediction plot after including the next design point from lower confidence bound exploration.
  - One-paragraph comparison of the different strategies.

- Your code.