Numerical Optimization 16: Surrogate Optimization

Qiang Zhu

University of Nevada Las Vegas

May 20, 2020

Overview

Prediction-Based Exploration

Gaussian process probes the probability distributions over the true objective function. These distributions can be used to guide an optimization process toward better design points. In prediction-based exploration, we select the minimizer of the surrogate function. If we use a Gaussian process surrogate model, prediction-based optimization has us select the minimizer of the mean function

$$\boldsymbol{x}^{m+1} = \underset{\boldsymbol{x} \in \chi}{\arg\min} \hat{\boldsymbol{\mu}}(\boldsymbol{x})$$

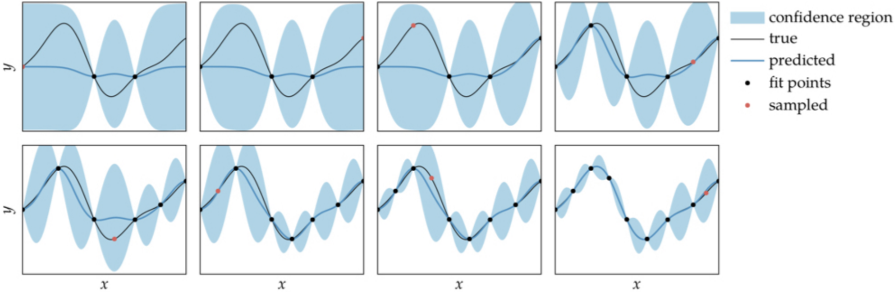
where $\hat{\boldsymbol{\mu}}(\boldsymbol{x})$ is the predicted mean of a Gaussian process at a design point x based on the previous m design points.
Prediction-based optimization does not take uncertainty into account, and new samples can be generated very close to existing samples, which is a waste of time.

Error-Based Exploration

Error-based exploration seeks to increase confidence in the true function. A Gaussian process can tell us both the mean and standard deviation at every point. The next sample point is:

$$\boldsymbol{x}^{m+1} = \underset{\boldsymbol{x} \in \mathcal{X}}{\arg\max} \hat{\boldsymbol{\sigma}}(\boldsymbol{x})$$

where $\hat{\boldsymbol{\sigma}}(\boldsymbol{x})$ is the predicted standard variance of a Gaussian process at a design point x based on the previous m design points.

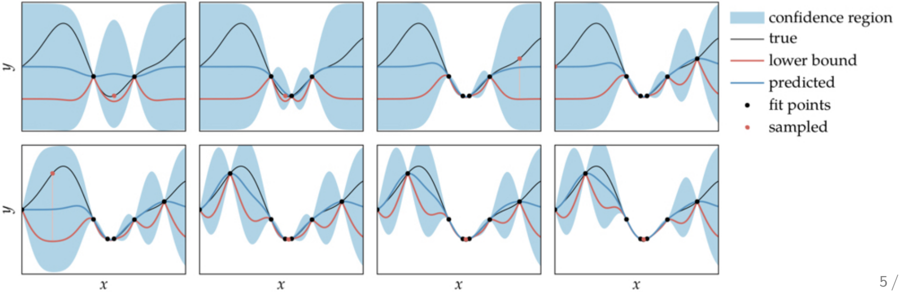


Lower Confidence Bound Exploration

The error-based exploration may sample the regions that are unpromising. Lower confidence bound exploration trades off between greedy minimization employed by prediction-based optimization and uncertainty reduction employed by error-based exploration with the following strategy,

$$LB(x) = \hat{\mu}(\boldsymbol{x}) - \alpha\hat{\sigma}(\boldsymbol{x})$$

$\alpha \geq 0$ is to control the trade-off between exploration and exploitation.



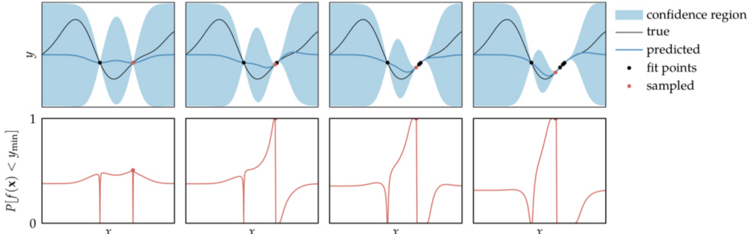
Probability of Improvement Exploration

We select the design point that maximizes the chance that the new point will be better than any other. The improvement for a function sampled at x producing $y = f(x)$ is

$$I(y) = \begin{cases} y_{\text{min}} - y & \text{if } y < y_{\text{min}} \\ 0 & \text{otherwise} \end{cases}$$

The probability of improvement at points where $\hat{\sigma} > 0$ is

$$P(y < y_{\text{min}}) = \int_0^{y_{\text{min}}} \mathcal{N}(y|\hat{\mu}, \hat{\sigma})dy = \Phi\left(\frac{y_{\text{min}} - \hat{\mu}}{\hat{\sigma}}\right)$$



Expected Improvement Exploration

We can also focus our exploration of points that maximize our expected improvement over the current best function value. Through a substitution

$$z = \frac{y - \hat{\mu}}{\hat{\sigma}} \qquad y`_{\text{min}} = \frac{y_{\text{min}} - \hat{\mu}}{\hat{\sigma}}$$

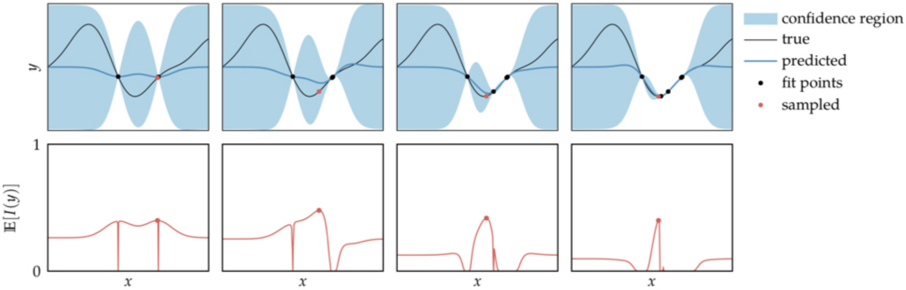
we can write the improvement as

$$I(y) = \begin{cases} \hat{\sigma}(y`_{\text{min}} - z) & \text{if } z < y`_{\text{min}} \text{ and } \hat{\sigma} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the predicted mean and standard deviation.
We can calculate the expected improvement using Gaussian process:

$$\begin{aligned} \mathbb{E}[I(y)] &= \hat{\sigma} \int_{-\infty}^{y`_{\text{min}}} \mathcal{N}(z|0,1)dz \\ &= (y_{\text{min}} - \hat{\mu})P(y \leq y_{\text{min}}) + \hat{\sigma}\mathcal{N}(y_{\text{min}}|\hat{\mu}, \hat{\sigma}2) \end{aligned}$$

Expected Improvement Exploration



Summary

- Gaussian processes can be used to guide the optimization process using a variety of strategies that use estimates of quantities such as the lower confidence bound, probability of improvement, and expected improvement.
- Some problems do not allow for the evaluation of unsafe designs, in which case we can use safe exploration strategies that rely on Gaussian processes.