

# Active Learning of Global Metamodels

Ross Alexander

*Department of Aeronautics and Astronautics*

*Stanford University*

Stanford, CA 94305

rbalexan@stanford.edu

## I. PROJECT UPDATE

### A. Progress So Far

Previously, we proposed to explore neural networks (NNs), Gaussian processes (GPs), and their deep variants (DNNs and DGPs) in global metamodeling tasks with active learning. Global metamodeling is the discipline concerned with constructing a surrogate model over the input space that is optimal with respect to the number of training examples. We employ active learning in global metamodeling (also called adaptive sampling) to improve sample efficiency and thus, optimality of the metamodel.

So far, we have implemented variance-based adaptive sampling for GPs on several 1D and 2D functions. First, we generate a small initial set of design points and objective function evaluations (e.g. a dataset  $\mathcal{D} = \{(x^{(i)}, y^{(i)})\}$ ) using a random sampling approach. We provide the initial sample set and fit a GP to the samples. We then evaluate the GP over a uniform grid to determine the next sample location according to the variance-based exploration strategy. Following this, we continually add the requested sample, re-fit the GP, and identify a new sample.

Figure 1 shows five iterations of active learning on the 1D  $\text{sinc}(x)$  function that was initialized with 12 random samples over the domain. We can see that the variance-based method performs well and quickly learns a good metamodel. Similarly, we show a single iteration of active learning on 2D Franke function in Figure 2. We are observing the cubic evaluation complexity, so extensions to higher dimensions is currently limited.

Some new additional directions to explore include generating the initial dataset using a central composite design, which could improve performance by reducing adaptive samples requested on the design space boundaries. Additionally, we may formulate a method to select samples based on expected improvement in integrated variance.

### B. Work Ahead

We should be able to at least complete a comparison between GPs and NNs, between GPs and DGPs, or between GPs with various adaptive sampling approaches. This will require developing the NN or DGP code along with coding up the calculation of performance metrics (i.e. RMSE, ISE, IV, and model generation time). We should have some results for these metamodels and active learning techniques on several general optimization benchmark functions of varying dimensionality and modality.

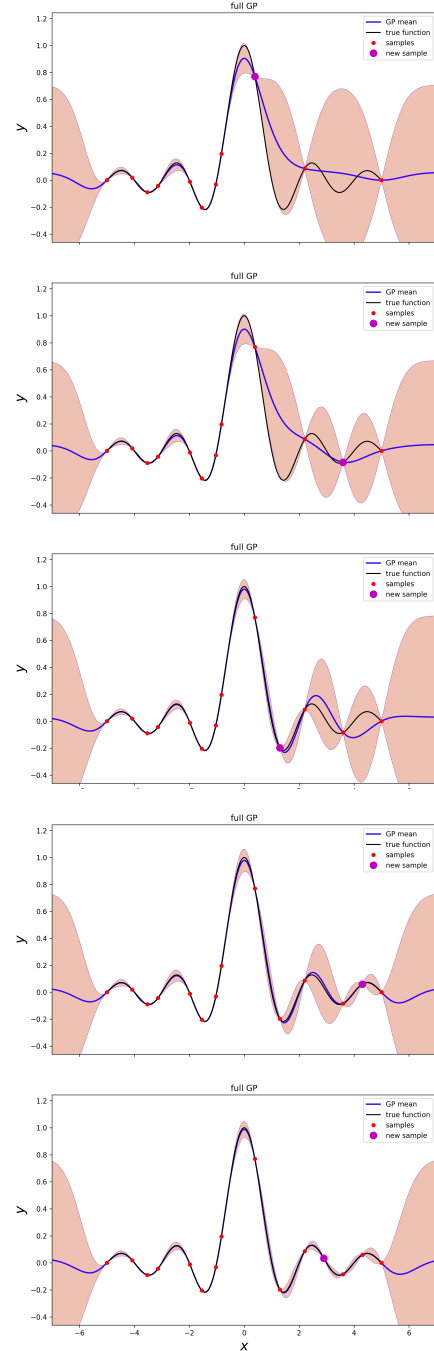


Fig. 1. Five iterations of variance-based active learning for a Gaussian process on the one-dimensional cardinal sine ( $\text{sinc}(x)$ ) function on  $x \in [-5, 5]$  initialized with 12 random samples.

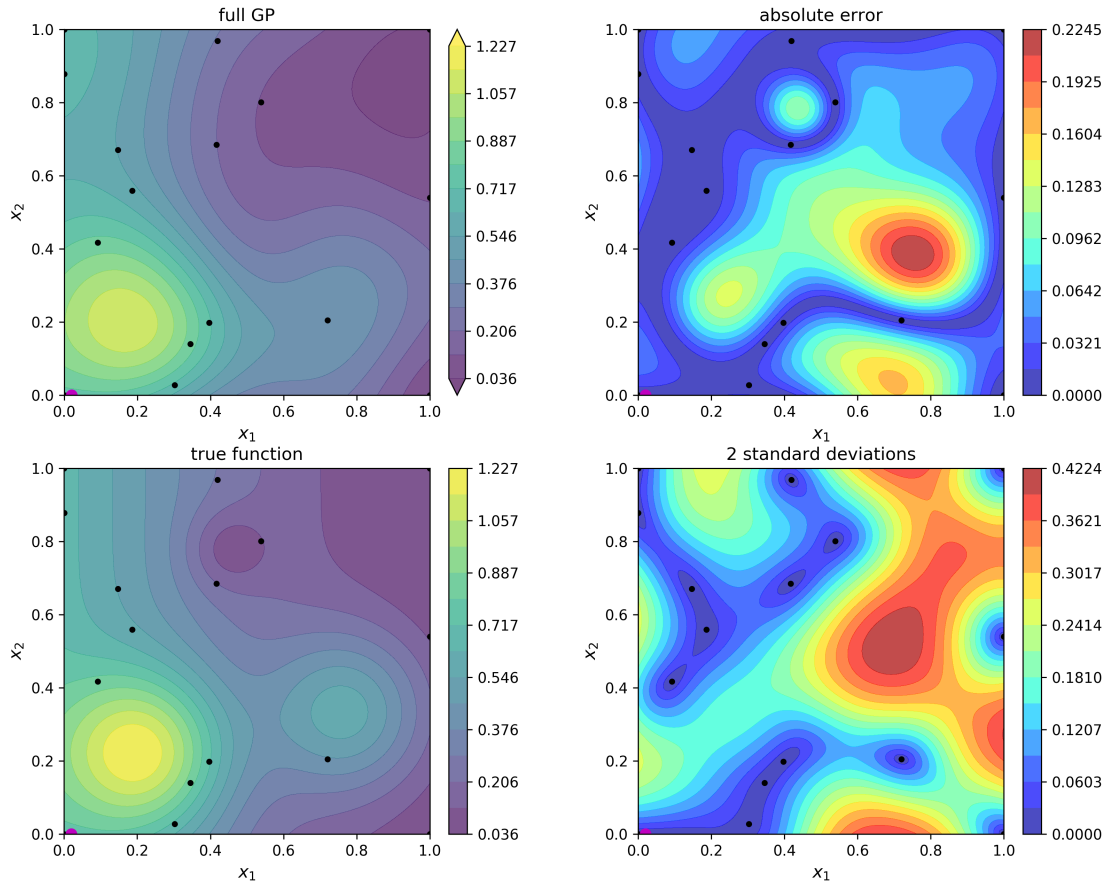


Fig. 2. A single iteration of variance-based active learning for a Gaussian process on the two-dimensional Franke function. The left plots show the predicted mean function and the true function. The right plots show the absolute error between the predicted mean function and the true function and magnitude of the two-sigma estimator from the GP.