ABSTRACT

In the standard setting of Bayesian optimization, we are tasked with the optimization of a black box function that is assumed to be highly expensive to evaluate. However, in many cases, cheap approximations may be available, such as computer simulations or training on smaller datasets. Multi-fidelity Bayesian optimization enables a flexible tradeoff between cost and accuracy by using these approximations to identify and eliminate low value regions cheaply while using expensive evaluations to exploit small promising regions with reduced uncertainty. Multi-fidelity strategies can achieve better regret bounds than single-fidelity methods that must rely on the objective function to explore the entire function space. We formalize the theoretical framework for the multi-fidelity setting and discuss techniques to approach the challenge of capturing shared latent relationships between mutually dependent fidelities, instead of effectively learning an independent surrogate for each. Recently, multi-output Gaussian processes have been implemented to model the joint distribution, but to avoid intractable computation of the acquisition function, imposed correlation structures include kernel convolution, strict correspondence between fidelity and cost, and hierarchical partitioning of the domain. We present the advantages of information-theoretic criteria for sequential querying and demonstrate empirical performance on both synthetic datasets and real-world applications.