An Improved Bayesian Framework for Quadrature

Henry Chai, Roman Garnett

Department of Computer Science and Engineering; Washington University in Saint Louis, Saint Louis, MO 63130, USA

Introduction

- Problem: estimating intractable integrals of the form $Z = \int f(x)\pi(x)\,\mathrm{d}x$

- Common techniques rely on Monte Carlo estimators, which are agnostic to any a priori knowledge.

- We offer a general framework for performing quadrature involving a broad class of constrained integrands, develop an instantiation of this framework that addresses some shortcomings of previous work and develop a novel procedure for tuning hyperparameters.

A framework for quadrature with constrained integrands

Let $Z = \int f(x)\pi(x)\,\mathrm{d}x$ be an intractable integral and let $f(x)$ be a constrained function where the constraint can be described as a warping of $\mathbb{R}$.

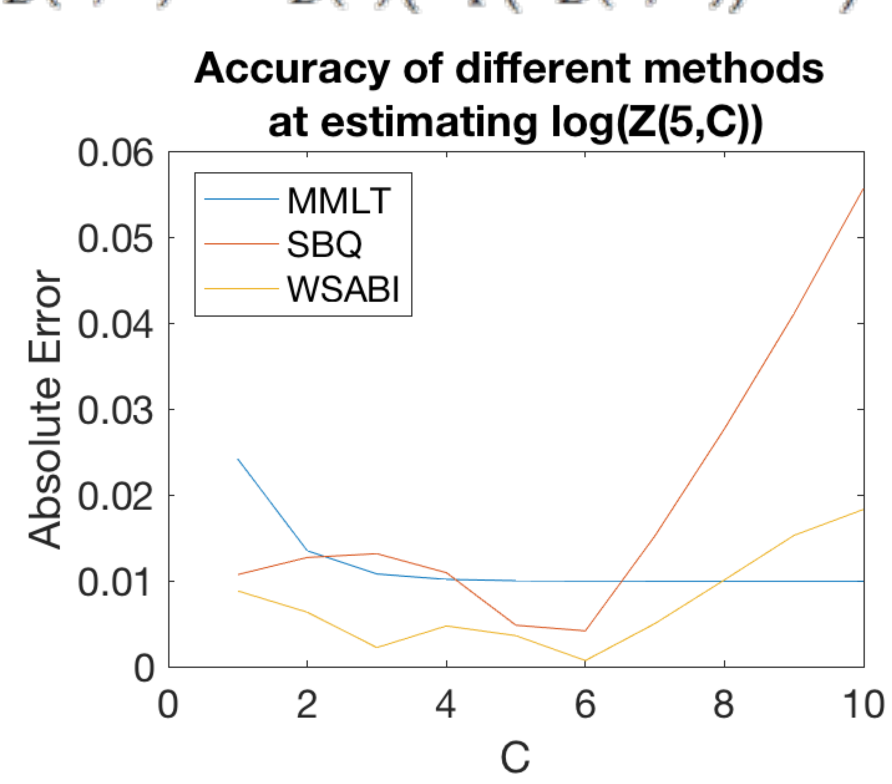
1. Determine a warping function ξ such that ξ maps from $\mathbb{R}$ to the range of f. Let $g(x) = \xi^{-1}(f(x))$. Place a GP prior on $g \sim \mathcal{GP}(\mu, \Sigma)$.
2. Calculate the mean and variance of the resulting probabilistic belief on f (using the first and second moments): $m(\mu(x), \Sigma(x,x))$ and $K(\mu(x), \mu(x'), \Sigma(x,x'))$. If these moments are not available in closed form (due to one's choice of ξ), they can approximated using Taylor series expansions. Approximate the probabilistic belief on f by a GP: $f \sim \mathcal{GP}(m, K)$.
3. Iterate until the budget of evaluations is expended:
   (a) Select a location to sample using uncertainty sampling, which selects the point that maximizes the posterior variance of f: $x^* = \arg\max_x K_D(x,x)$, and observe $g(x^*)$.
   (b) Update the posterior belief $g \sim \mathcal{GP}(\mu_D, \Sigma_D)$, fitting the hyperparameters associated with the GP on g by maximizing the marginal likelihood of the observations of f, using the approximate GP belief on f.
4. Calculate the belief about the value of the integral given all observations D: $Z|D \sim N(\int m_D(x)\pi(x)\,\mathrm{d}x, \iint K_D(x,x')\pi(x)\pi(x')\,\mathrm{d}x\,\mathrm{d}x')$. Again, if these quantities cannot be computed in closed form, they can be estimated using Taylor series expansions. This follows from Isserlis' theorem and the fact that for most common choices of μ and Σ, $\mu(x)^a\Sigma(x,x')^b\mu(x')^c$ can be integrated exactly for non-negative a, b and c.

Moment-matched log transformation

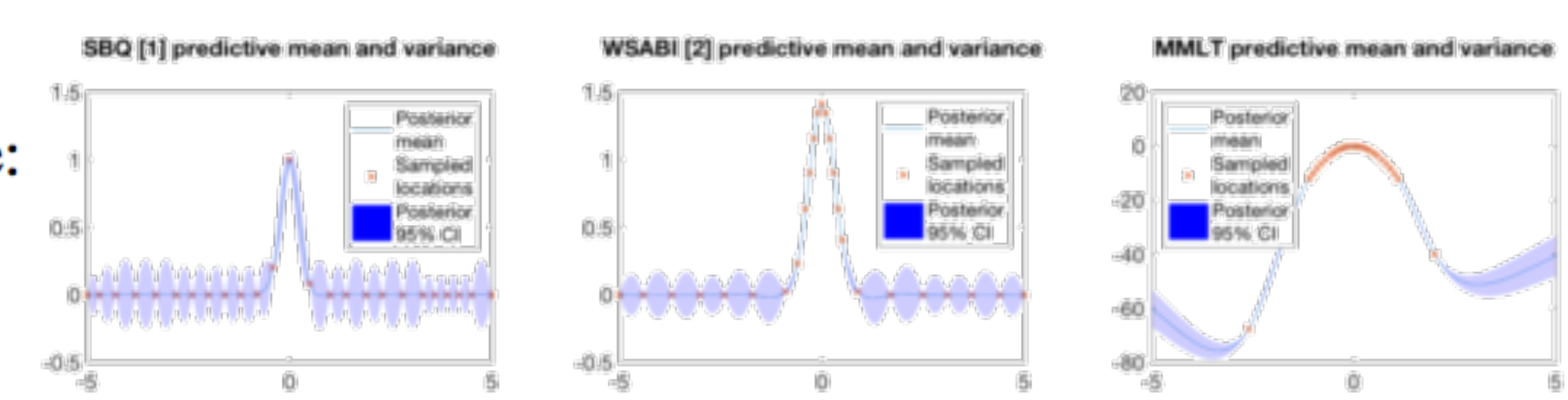
Suppose f is constrained to be non-negative; an appropriate choice for ξ in this setting is log. If $g \sim \mathcal{GP}(\mu_D(x), \Sigma_D(x,x'))$, then the mean and variance of the true posterior of f are:

$$m_D(x) = \exp(\mu_D(x) + 1/2\Sigma_D(x,x'))$$

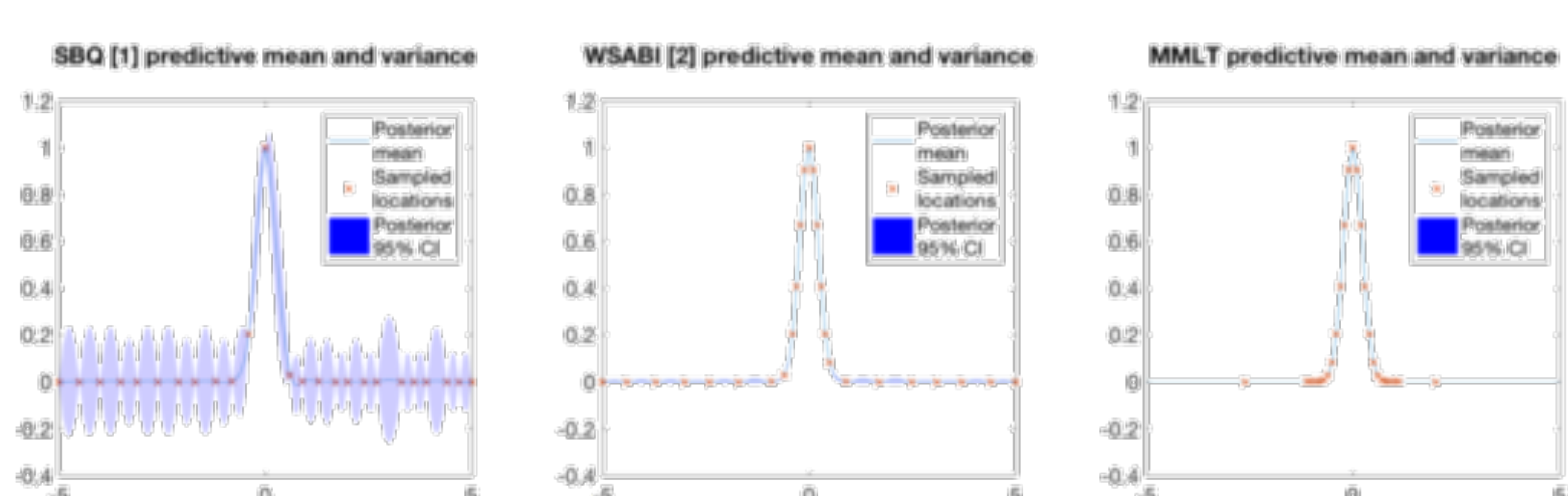
$$\text{and } K_D(x,x') = m_D(x)(\exp(\Sigma_D(x,x')) - 1)m_D(x')$$



Toy problem: $Z(a,C) = \int_{-a}^{a} \exp(-Cx^2/2)(1/2a)\,\mathrm{d}x$



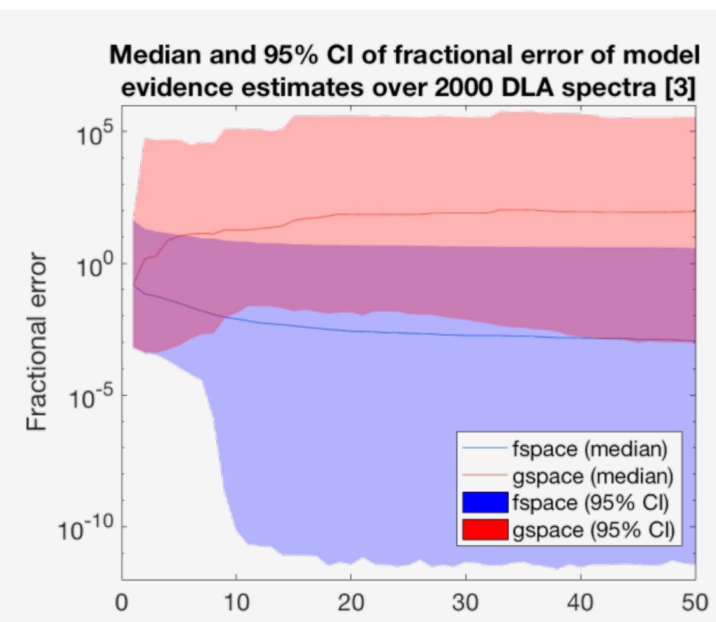
Posterior belief of g for $Z(5,10)$

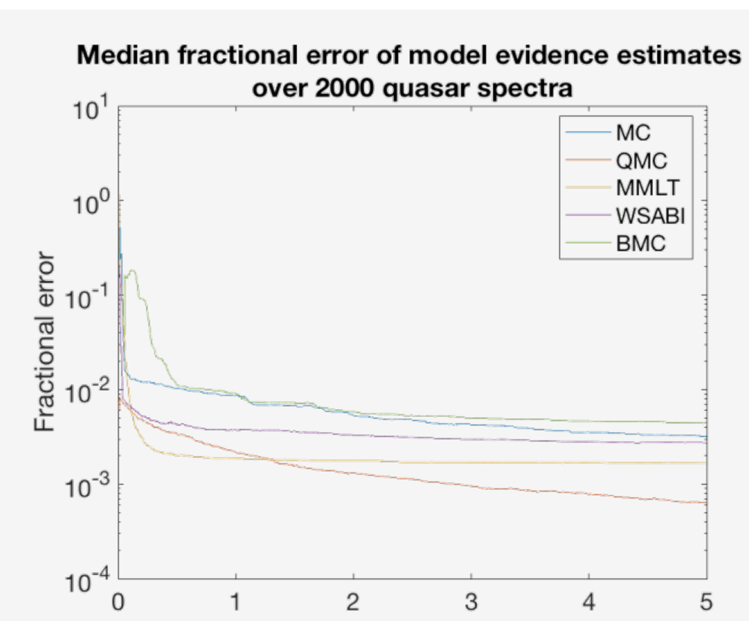


Posterior belief of f for $Z(5,10)$

Hyperparameter optimization

Instead of setting the hyperparameters of the GP on g by maximizing the marginal likelihood of $D = \{x, g(x)\}$, we maximize the marginal likelihood of $D = \{x, f(x)\}$ using the approximate GP belief on f.





References

[1] C. E. Rasmussen and Z. Ghahramani. Bayesian Monte Carlo. NIPS, 2003.
[2] T. Gunter, M. A. Osborne, R. Garnett, P. Hennig and S.J. Roberts. Sampling for inference in probabilistic models with fast Bayesian quadrature. In NIPS, 2014.
[3] R. Garnett, S. Ho, S. Bird, J. Schneider. Detecting Damped Lyman-α Absorbers with Gaussian Processes. In *Monthly Notices of the Royal Astronomical Society*, 2016.

Washington University in St.Louis

MISSOURI EPSCoR

MISSOURI TRANSECT

NSF