Automated Model Selection Using Bayesian Qaudarture

Henry Chai[1] Jean-François Ton[2] Michael A. Osborne[2] Roman Garnett[1]

[1]Washington University in St Louis, [2]University of Oxford

UNIVERSITY OF OXFORD

DEPARTMENT OF STATISTICS

Washington University in St. Louis

ICML 2019

Posterior Probability of a Model

The posterior probability of model i (PPM) is defined as

$$z_i = Z_i / \sum_j Z_j$$

where

$$Z_i = \int f(\mathcal{D}|\theta_i)\pi(\theta_i)d\theta_i$$

is the model evidence for model i and θ_i are the parameters of model M_i, $f(\mathcal{D}|\theta_i)$ is the likelihood and $\pi(\theta_i)$ is the prior.

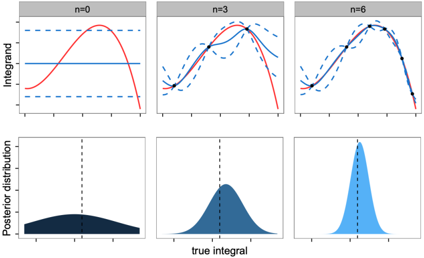
Motivation

Comparing the model evidences usually relies on Monte Carlo (MC)methods, which converge slowly and are unreliable for expensive models.

Bayesian Quadrature

Given some intractable integral

$$Z = \int f(\theta)\pi(\theta)d\theta$$



[1]

The goal is it to choose the points where we evaluate the GP efficiently. For example it does not make sense to take 2 points that are really "close".

[1] https://warwick.ac.uk/girolami

Motivation

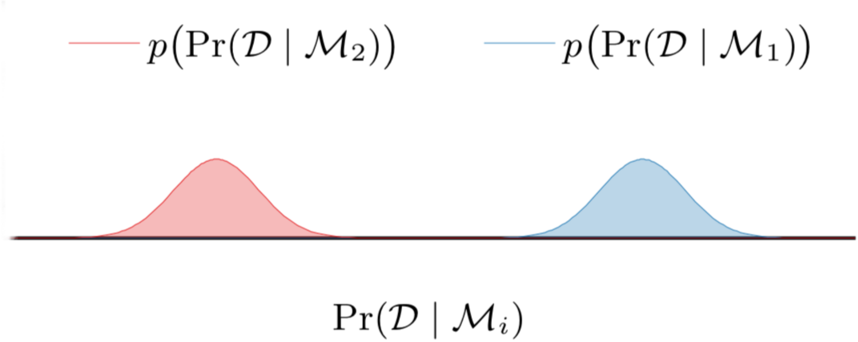
Hence we can apply Bayesian Quadrature to estimated the model evidences

Motivation

Hence we can apply Bayesian Quadrature to estimated the model evidences

However, this may waste computation by producing an overly-accurate estimate for the evidence of a clearly poor model

Motivation



$p(\Pr(\mathcal{D} \mid \mathcal{M}_2))$
$p(\Pr(\mathcal{D} \mid \mathcal{M}_1))$
$\Pr(\mathcal{D} \mid \mathcal{M}_i)$

Idea

We propose a **sampling algorithm** which targets the most important quantity; **Posterior Probability of the Model** i.e. $z_i = Z_i / \sum_j Z_j$.

Idea

We propose a **sampling algorithm** which targets the most important quantity; **Posterior Probability of the Model** i.e. $z_i = Z_i / \sum_j Z_j$.

NOTE that our task is to decide where to evaluate the likelihood f

$$Z_i = \int f(\mathcal{D}|\theta_i)\pi(\theta_i)d\theta_i$$

Idea

We propose a **sampling algorithm** which targets the most important quantity; **Posterior Probability of the Model** i.e. $z_i = Z_i / \sum_j Z_j$.

NOTE that our task is to decide where to evaluate the likelihood f

$$Z_i = \int f(\mathcal{D}|\theta_i)\pi(\theta_i)d\theta_i$$

Hence we utilize the following Mutual Information

$$MI(z_i, f(\mathcal{D}|\theta_i))$$

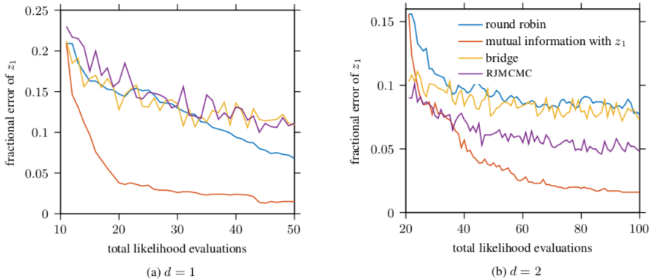
yielding efficient acquisition of samples across disparate model spaces when likelihood observations are limited.

Synthetic dataset

Model selection task between two zero-mean GP models

- One with RBF kernel
- Other with Matern5/2

The observed dataset $\mathcal{D}$ consists of $5d$ observations from a d-dimensional, zero-mean GP with a RBF kernel



End

Thanks you for your attention
If interested visit poster 06:30 – 09:00 PM @ Pacific Ballroom

#**219**