BAYESIAN STATISTICS PROJECT

FRANK-WOLFE BAYESIAN QUADRATURE: PROBABILISTIC INTEGRATION WITH THEORETICAL GUARANTEES

AUTHORS

LYN BOUSSENGUI, ANTOINE BROWAEYS
ACHRAF BZILI, NICOLAS CLOAREC

Notations

We start with some notations we will use along this report.

- $p(\cdot)$ denotes the function $p : x \mapsto p(x)$.

Introduction

The goal of this article [1] is to compute efficiently the integrals of the form $\displaystyle\int_{\mathcal{X}} f(x)p(x)\mathrm{d}x$ where $\mathcal{X} \subseteq \mathbb{R}^d$ is a measurable space, $d \geq 1$ integer representing the dimension of the problem, p a probability density with respect to the Lebesgue measure on $\mathcal{X}$ and $f : \mathcal{X} \to \mathbb{R}$ is a *test*-function.

We will use the common approximation

$$\int_{\mathcal{X}} f(x)p(x)\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i) \tag{1}$$

but of course the real challenge lies in the choice of sequences $\{x_i\}$ and $\{w_i\}$:

- **Monte Carlo**: $w_i = \frac{1}{n}$ and x_i realization of multivariate random variable $X_i \overset{iid}{\sim} X$ where X has $p(\cdot)$ as probability distribution.
- **Kernel herding**:
- **Quasi-Monte Carlo**:

In the **Frank-Wolfe Bayesian Quadrature**, we have

☞ $\{w_i\}$ which appear naturally in the Bayesian Quadrature by taking the expectation of a posterior distribution (described in section 2),

☞ $\{x_i\}$ selected by the Frank-Wolfe algorithm in order to minimize a posterior variance (described in section 3).

The main interest of the method developed in [1] is the super fast *exponential* convergence to the true value of the integral compared to the other methods mentioned above.

1 Theoretical framework

2 Bayesian Quadrature

3 Frank-Wolfe algorithm

References

[1] François-Xavier Briol, Chris J. Oates, Mark Girolami, and Michael A. Osborne. Frank-wolfe bayesian quadrature: Probabilistic integration with theoretical guarantees.