Kernel Methods, Quadrature or Sampling, and Probabilistic Numerics

Motonobu Kanagawa

Prob Num 2018
London, April 2018

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Max Planck Institute for
Intelligent Systems

erc
some of the presented work is supported by the European Research Council.

Short Self-Introduction

- Research Scientist (Postdoc) at the PN group in Max Planck
  - September 2017 $\sim$
  - Working with Dr. Philipp Hennig
- PhD student and Postdoc at the Institute of Statistical Mathematics, Tokyo
  - $\sim$ August 2017
  - Worked with Prof. Kenji Fukumizu

Kernel Methods

- Kernel mean embedding of distributions

$$\mu_P = \int k(\cdot, x)dP(x)$$

Kernel Methods

- Kernel mean embedding of distributions

$$\mu_P = \int k(\cdot, x) dP(x)$$

- Kernel Monte Carlo Filter [Kanagawa et al., 2016a]: Combination of
  1. Kernel Bayes' Rule [Fukumizu et al., 2013]
  2. Monte Carlo sampling
  3. Kernel Herding [Chen et al., 2010]

- Kernel mean embedding of distributions

$$\mu_P = \int k(\cdot, x) dP(x)$$

- Kernel Monte Carlo Filter [Kanagawa et al., 2016a]: Combination of
  1. Kernel Bayes' Rule [Fukumizu et al., 2013]
  2. Monte Carlo sampling
  3. Kernel Herding [Chen et al., 2010]
- Getting interested in Kernel Herding...
  - Greedy approach to deterministic sampling or quadrature

$$x_t := \arg\min_{x \in \mathcal{X}} \left\| \mu_P - \frac{1}{t} \sum_{i=1}^{t-1} k(\cdot, x_i) - \frac{1}{t} k(\cdot, x) \right\|_{\mathcal{H}_k}$$

Quadrature or Sampling with Kernels

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

- Convergence analysis for kernel-based quadrature rules in misspecified settings [Kanagawa et al., 2016b, Kanagawa et al., 2017]

Quadrature or Sampling with Kernels

- Convergence analysis for kernel-based quadrature rules in misspecified settings [Kanagawa et al., 2016b, Kanagawa et al., 2017]
- Given $(w_i, x_i)_{i=1}^n \subset \mathbb{R} \times \mathcal{X}$ such that

$$\lim_{n \to \infty} \left\| \mu_P - \sum_{i=1}^{n} w_i k(\cdot, x_i) \right\|_{\mathcal{H}_k} = 0,$$

where $\mathcal{H}_k$ is the RKHS of k, what can we say about the error

$$\left| \int f(x) dP(x) - \sum_{i=1}^{n} w_i f(x_i) \right|,$$

for a misspecified $f \notin \mathcal{H}_k$?

Probabilistic Numerics

Current research interests include ...

Current research interests include ...

1. Connections between Gaussian processes and kernel methods
   - Currently writing a review paper ... to be submitted soon.

Probabilistic Numerics

Current research interests include ...

1. Connections between Gaussian processes and kernel methods
   - Currently writing a review paper ... to be submitted soon.
2. Decision theoretic viewpoint for probabilistic numerics
   - Why uncertainty matters? Because we need to make decisions based on numerics! (Talk at SIAM-UQ)

Current research interests include ...

1. Connections between Gaussian processes and kernel methods
   - Currently writing a review paper ... to be submitted soon.
2. Decision theoretic viewpoint for probabilistic numerics
   - Why uncertainty matters? Because we need to make decisions based on numerics! (Talk at SIAM-UQ)
3. Bayesian quadrature
   - Transformation of a Gaussian process prior (e.g., WSABI)
   - High-dimensional integration
   - Stein's method for an unnormalized density [Oates et al., 2017].

Probabilistic Numerics

Current research interests include ...

1. Connections between Gaussian processes and kernel methods
   - Currently writing a review paper ... to be submitted soon.
2. Decision theoretic viewpoint for probabilistic numerics
   - Why uncertainty matters? Because we need to make decisions based on numerics! (Talk at SIAM-UQ)
3. Bayesian quadrature
   - Transformation of a Gaussian process prior (e.g., WSABI)
   - High-dimensional integration
   - Stein's method for an unnormalized density [Oates et al., 2017].
4. (Approximate Bayesian Computation)
   - Application of Kernel herding [Kajihara et al., 2018]

- Chen, Y., Welling, M., and Smola, A. (2010).
  Supersamples from kernel-herding.
  In *Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence (UAI 2010)*, pages 109–116.

- Fukumizu, K., Song, L., and Gretton, A. (2013).
  Kernel Bayes' rule: Bayesian inference with positive definite kernels.
  Journal of Machine Learning Research, 14:3753–3783.

- Kajihara, T., Yamazaki, K., Kanagawa, M., and Fukumizu, K. (2018).
  Kernel recursive ABC: Point estimation with intractable likelihood.
  ArXiv e-prints, stat.ML 1802.08404.

- Kanagawa, M., Nishiyama, Y., Gretton, A., and Fukumizu, K. (2016a).
  Filtering with state-observation examples via kernel monte carlo filter.
  Neural Computation, 28(2):382–444.

- Kanagawa, M., Sriperumbudur, B. K., and Fukumizu, K. (2016b).

Convergence guarantees for kernel-based quadrature rules in misspecified settings.

In *Advances in Neural Information Processing Systems 29*.

- Kanagawa, M., Sriperumbudur, B. K., and Fukumizu, K. (2017).
  Convergence Analysis of Deterministic Kernel-Based Quadrature Rules in Misspecified Settings.
  arXiv:1709.00147 [cs, math, stat].
  arXiv: 1709.00147.
- Oates, C. J., Girolami, M., and Chopin, N. (2017).
  Control functionals for Monte Carlo integration.
  Journal of the Royal Statistical Society, Series B, 79(2):323–380.