Modern Gaussian Processes: Scalable Inference and Novel Applications

(Part IV) Theory & Code

Edwin V. Bonilla and **Maurizio Filippone**

CSIRO's Data61, Sydney, Australia and EURECOM, Sophia Antipolis, France

July 14th, 2019

Outline

1. Theory for GPs
   - Asymptotics & Consistency
   - GPs & Stochastic Differential Equations
   - Other Interesting Topics
2. Code

Theory for GPs

Asymptotics & Consistency

- The GP posterior mean minimizes the following functional:

$$J(f) = \frac{1}{2}\|f\|^2_{\mathcal{H}} + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(\mathbf{x}_i))$$

  where $\|f\|^2_{\mathcal{H}}$ is the RKHS norm corresponding to the covariance function κ.

- What happens when $N \to \infty$?

Asymptotics & Consistency

- The GP posterior mean minimizes the following functional:

$$J(f) = \frac{1}{2}\|f\|^2_{\mathcal{H}} + \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(\mathbf{x}_i))$$

  where $\|f\|^2_{\mathcal{H}}$ is the RKHS norm corresponding to the covariance function κ.
- What happens when $N \to \infty$?
- f converges to $\mathbb{E}_{p(y,\mathbf{x})}[y|\mathbf{x}]$ ...
- ... under some regularity conditions (nondegenerate κ, regression function well-behaved)

GPs & Stochastic Differential Equations

- Consider the Markov process:

$$a_m \frac{d^m f(x)}{dx^m} + a_{m-1} \frac{d^{m-1} f(x)}{dx^{m-1}} + \dots a_1 \frac{df(x)}{dx} + a_0 f(x) = w(x)$$

  where $w(x)$ is a zero-mean white-noise process.
- The solution is a GP
- The covariance depends on the form of the SDE
- Solving SDEs is easy in low dimensions!
- We can solve GPs in $\mathcal{O}(N \log N)$

Saatçi, *Ph.D. Thesis*, 2011

Other Interesting Topics

- Average-case Learning Curves
- PAC-Bayesian Analysis
- Theory for Sparse GPs - Best Paper Award ICML 2019

Code

Code for Gaussian Processes

- python
  - GPy
- MatLab
  - gptoolbox
- R
  - kernlab

Code for Gaussian Processes - With Automatic Differentiation

- TensorFlow:
  - GPflow
  - AutoGP
- PyTorch
  - CandleGP

Deep Gaussian Processes

- TensorFlow:
  - GPflow
  - Doubly-Stochastic DGPs
- PyTorch
  - DGPs with Random Features
- Theano
  - DGPs with Inducing Points & Exp. Propagation