Data Analysis and Machine Learning: Machine learning with Gaussian Processes

Christian Forssén[1] Morten Hjorth-Jensen[2,3]

Department of Physics, Chalmers University of Technology, Sweden[1]

Department of Physics, University of Oslo[2]

Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University[3]

Mar 19, 2018

© 1999-2018, Morten Hjorth-Jensen. Released under CC Attribution-NonCommercial 4.0 license

What is a Gaussian Process?

- We have considered splines and kernel regression methods. These

require choice of somewhat arbitrary set of knots.

- Antoher possibility is to setup a prior distribution for the regression function using a *Gaussian Process*.
- This is a very flexible class of models that has distinct computational and theoretical advantages. It can be viewed as a potentially infinite-dimensional generalization of Gaussian distributions.
- See the excellent (and free) book Gaussian Processes for Machine Learning by Carl Edward Rasmussen and Christopher K. I. Williams.

Gaussian process regression

- Realizations from a Gaussian process correspond to random functions
- Let us first consider an unknown regression function $\mu(x)$ that depends on a single, continuous variable x.
- The Gaussian process is written as $\mu \sim \mathrm{GP}(m, k)$, and is parametrized in terms of a mean function $m(x)$ and a covariance function $k(x, x')$.
- The GP prior on μ describes it as a random function for which the values at any set of N prespecified points $\{x_i\}_{i=1}^N$ are a draw from a N-dimensional normal distribution

$$\mu(x_1), \ldots \mu(x_N) \sim \mathrm{N}\left((m(x_1), \ldots, m(x_N)), K(x_1, \ldots, x_N)\right),$$

with mean m and covariance K.

Topics

- More matematical details
- The role of the covariance function (different kernels)
- multidimensional case
- examples.