

Lecture 21: Gaussian process regression

Professor Ilias Bilonis

Gaussian process regression - Overview

Regression

You are given n observations consisting of:

$$\mathbf{x}_{1:n} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$$

(inputs, features, ...)

$$y_{1:n} = \{y_1, \dots, y_n\}$$

(outputs, targets, labels, ...)

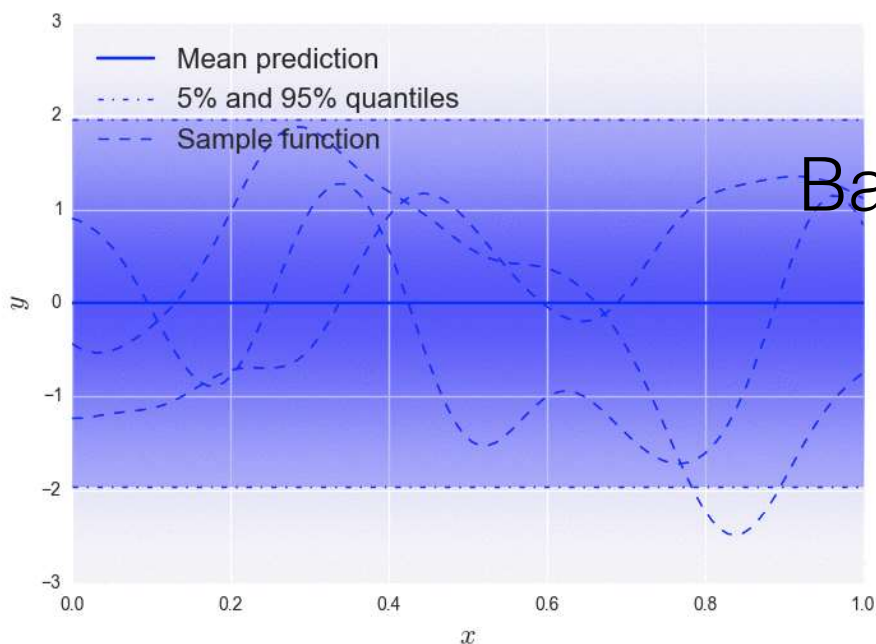
Continuous outputs

Problem: Use the data to learn the map between x and y

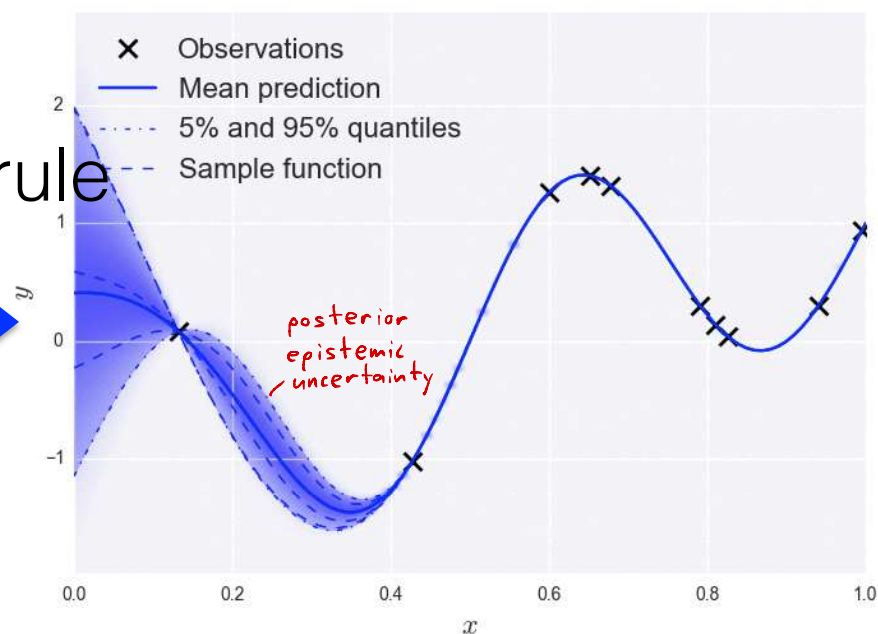
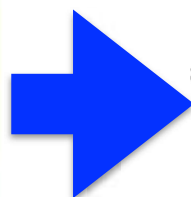
Gaussian process regression

- The “best” regression you can do when:
 - The dimensionality of the features is not too big (<100).
 - The number of observations is not too large (<5000).

How does Gaussian process regression work?



Bayes' rule



Prior GP

Posterior GP

assumptions
are
encoded

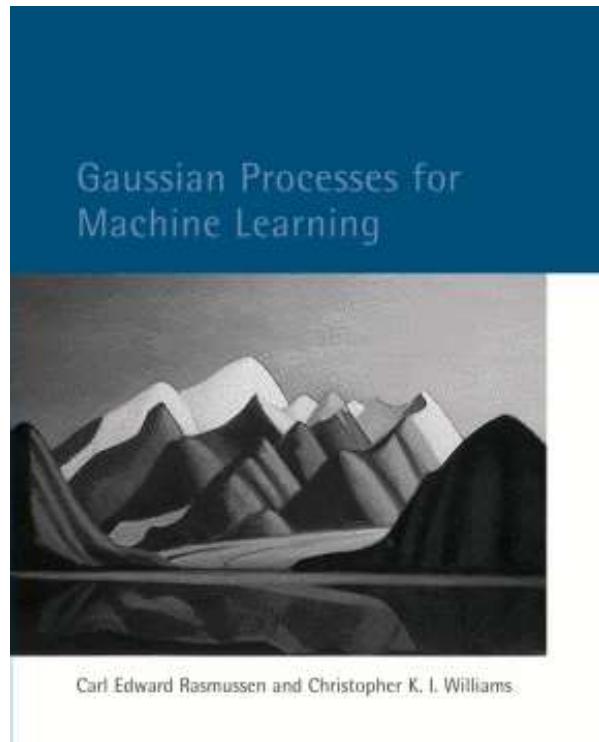
GP helps to assign probability measure
on a function space that maps from
features to targets

PREDICTIVE
SCIENCE LABORATORY

Why is the Gaussian process regression so powerful?

- It allows you to encode your prior beliefs about the regularity of the function (smoothness, length scale, variance). - in a non-parametric way
- It does not overfit because it is fully Bayesian.
- It is easy to put together from intuitive pieces.

The best book on the Subject



Gaussian Processes for Machine Learning
Carl Edward Rasmussen and Christopher K. I.
Williams
The MIT Press, 2006. ISBN 0-262-18253-X.

Free online at www.gaussianprocess.org.

The best code on the subject

GPy (in Python) from the group of N. Lawrence @ University of
Sheffield

<https://github.com/SheffieldML/GPy>