Lecture 21: Gaussian process regression

Professor Ilias Bilionis

Gaussian process regression - Overview

Regression

Your are given n observations consisting of:

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

(inputs, features, …)

$$\mathbf{y}_{1:n} = \{y_1, \ldots, y_n\}$$

Continuous outputs

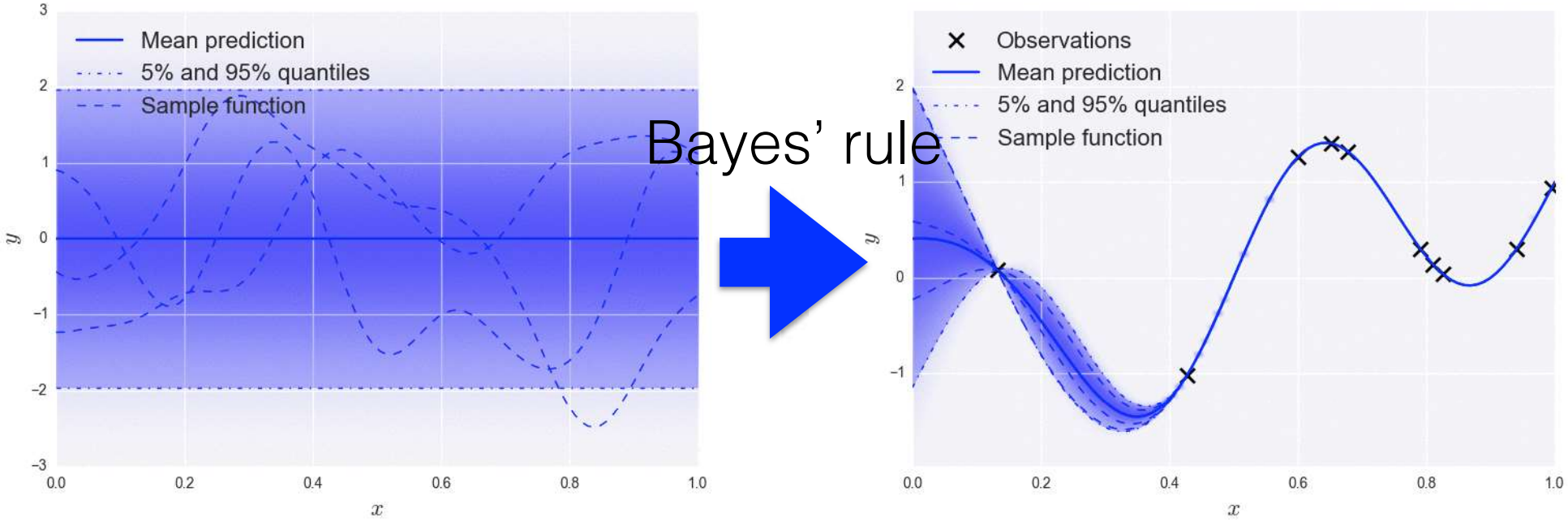
(outputs, targets, labels, …)

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?



Prior GP

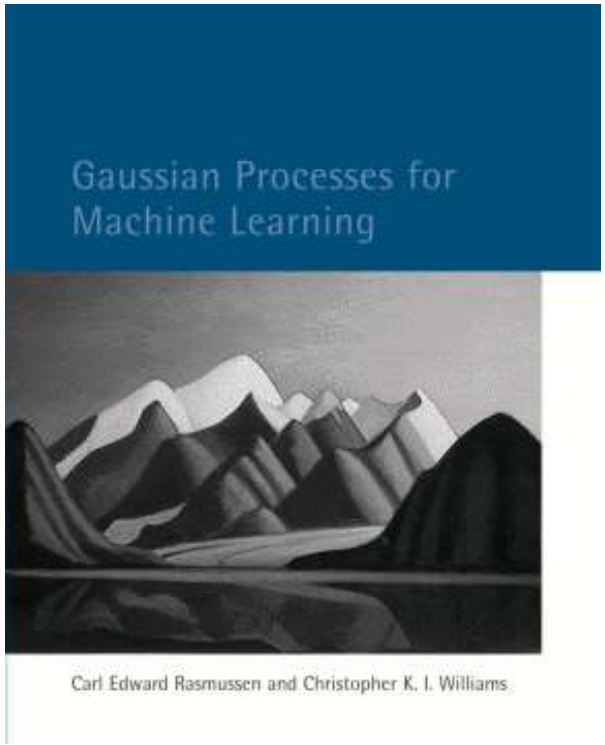
Posterior GP

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).
- It does not overfit because it is fully Bayesian.
- It is easy to put together from intuitive pieces.

The best book on the Subject

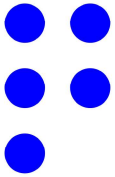


Gaussian Processes for Machine Learning
Carl Edward Rasmussen and Christopher K. I. Williams
The MIT Press, 2006. ISBN 0-262-18253-X.

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



PREDICTIVE
SCIENCE LABORATORY