Finite and infinite basis GPs

Carl Edward Rasmussen

October 13th, 2016

Key concepts

- Should we use finite or infinite models?
- GPs are a fancy way of using infinite models, but
  - will it actually make any difference in practise?
- finite models correspond to much stronger assumptions about the data
- therefore, we don't want to use finite models
- a GP with squared exponential covariance function corresponds to an infinite linear in the parameters model with Gaussian bumps everywhere
- illustrate the difference

Cromwell's dictum

I beseech you, in the bowels of Christ, consider it possible that you are mistaken

— Oliver Cromwell, 1650

From infinite linear models to Gaussian processes

Consider the class of functions (sums of squared exponentials):

$$f(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=-N/2}^{N/2} \gamma_n \exp(-(x - \tfrac{n}{\sqrt{N}})^2), \quad \text{where } \gamma_n \sim \mathcal{N}(0,1), \ \forall n$$

$$\stackrel{?}{=} \int_{-\infty}^{\infty} \gamma(u) \exp(-(x-u)^2)\, du, \quad \text{where } \gamma(u) \sim \mathcal{N}(0,1), \ \forall u.$$

The mean function is:

$$\mu(x) = E[f(x)] = \int_{-\infty}^{\infty} \exp(-(x-u)^2) \int_{-\infty}^{\infty} \gamma(u) p(\gamma(u)) d\gamma(u)\, du = 0,$$

and the covariance function:

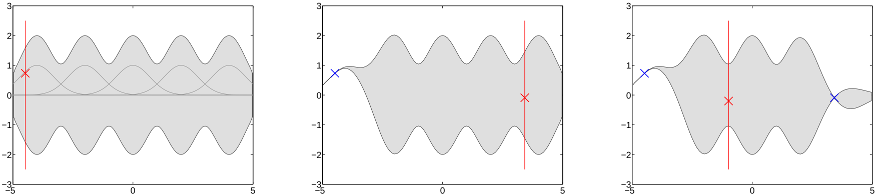
$$E[f(x)f(x')] = \int \exp\big(-(x-u)^2 - (x'-u)^2\big) du$$

$$= \int \exp\big(-2(u - \frac{x+x'}{2})^2 + \frac{(x+x')^2}{2} - x^2 - x'^2\big) du \stackrel{?}{\propto} \exp\big(-\frac{(x-x')^2}{2}\big).$$

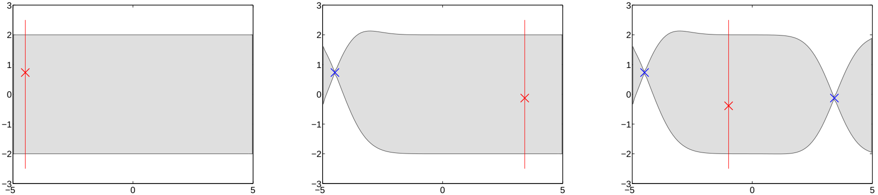
Thus, the squared exponential covariance function is equivalent to regression using infinitely many Gaussian shaped basis functions placed everywhere, not just at your training points!

Using finitely many basis functions may be dangerous!(1)

Finite linear model with 5 localized basis functions)



Gaussian process with infinitely many localized basis functions



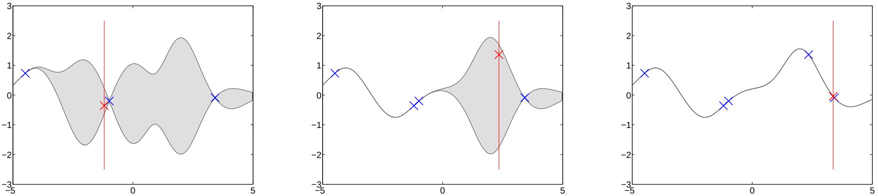
prior

posterior after observing one single datapoint

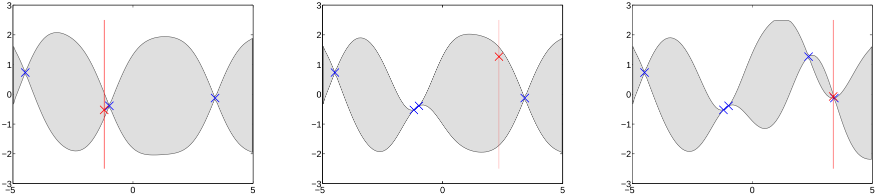
posterior after observing two datapoints.

Using finitely many basis functions may be dangerous!(2)

Finite linear model with 5 localized basis functions)

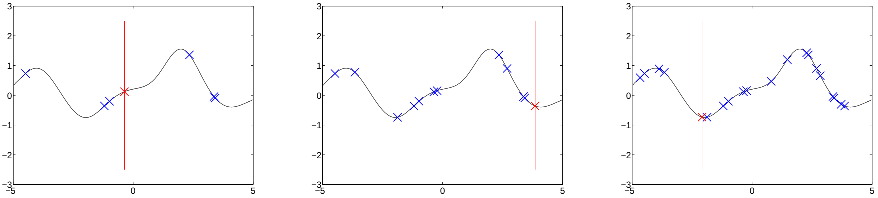


Gaussian process with infinitely many localized basis functions



Using finitely many basis functions may be dangerous!(3)

Finite linear model with 5 localized basis functions)



Gaussian process with infinitely many localized basis functions