LINKÖPING UNIVERSITY
Department of Computer and Information Science
Division of Statistics and Machine Learning
Mattias Villani

2015-11-15
Introduction to Machine Learning
Master and PhD course

Gaussian Processes - Computer Lab

Deadline: See LISAM
Teacher: Mattias Villani
Grades: Pass/Fail
Submission: Via LISAM

The lab is to be reported in a concise report.
You may use any programming language for the lab.
Attach your code in LISAM as separate files.

1. This exercise is concerned with the Gaussian Process Regression:

$$y = f(x) + \varepsilon \quad \varepsilon \sim N(0, \sigma_n^2)$$

Here we go:

(a) Write your own code for simulating from the posterior distribution of $f(x)$ using the squared exponential kernel. The function (name it posteriorGP) should return vectors with the posterior mean and variance of f, both evaluated at a set of x-values ($x^\star$). You can assume that the prior mean of f is zero for all x. The function should have the following inputs:

i. x (vector of training inputs)
ii. y (vector of training targets/outputs)
iii. xStar (vector of inputs where the posterior distribution is evaluated, i.e. $x^\star$. As in my slides).
iv. hyperParam (vector with two elements σ_f and ℓ)
v. sigmaNoise (σ_n).
[Hint: I would write a separate function for the Kernel (see my GaussianProcess.R function on the course web page) which is then called from the posteriorGP function.]

(b) Now let the prior hyperparameters be $\sigma_f = 1, \ell = 0.3$. Update this prior with a single observation: $(x, y) = (0.4, 0.719)$. Assume that the noise standard deviation is known to be $\sigma_n = 0.1$. Plot the posterior mean of f over the interval $x \in [-1, 1]$. Plot also 95% probability (pointwise) bands for f.

(c) Update your posterior from 1c) with another observation: $(x, y) = (-0.6, -0.044)$. Plot the posterior mean of f over the interval $x \in [-1, 1]$. Plot also 95% probability bands for f. [Hint: updating the posterior after one observation with a new observation gives the same result as updating the prior directly with the two observations. Bayes is beautiful!]

x	-1.0	-0.6	-0.2	0.4	0.8
y	0.768	-0.044	-0.940	0.719	-0.664

Table 1: Simple data set for GP regression.

(d) Compute the posterior distribution of f using all 5 data points in Table 1 below (note that the two previous observations are included in the table). Plot the posterior mean of f over the interval $x \in [-1, 1]$. Plot also 95% probability intervals for f.

(e) Repeat 1d), this time with the hyperparameters $\sigma_f = 1, \ell = 1$. Compare the results.

(f) Congratulations, you have created your own Gaussian process regression code! And you know *all* the little details in it. Sweet!

2. Try out your brand new code from the previous problem on the `JapanTemp.dat` data set (available in LISAM) with `time` as covariate and `temp` as response variable. The data set contains a year of daily temperatures for some place in Japan. Document the effect on the posterior from changing σ_n and the prior hyperparameters.

Good luck! Hopefully things will go smoothly (yes, pun intended).

- Mattias