Pattern Recognition: Assignment 8

Due on Tuesday, June 26 2012, 23:59

Bernhard X. Kausler (TA), Summer Term 2012

patternrecognition@hci.iwr.uni-heidelberg.de

`http://hci.iwr.uni-heidelberg.de/MIP/Teaching/pr/`

Gaussian Processes

In this assignment you will learn about Gaussian processes. First, we sample from a Gaussian process and investigate the influence of the covariance function and observed values. Subsequently, we apply Gaussian processes to regression and explore different interpretations of GP regression.

Prob. 1: Sampling From Gaussian Processes

In this problem we will sample from a Gaussian process and investigate the influence of different covariance functions and conditioning on observed values.
The values Y^* of finite set of n points drawn from a Gaussian process at locations $\mathbf{x}$ is distributed according to a multivariate Gaussian

$$p(Y^*) = \mathcal{N}(Y^*|\mathbf{0}, \mathbf{K}(\mathbf{x}))$$

with a covariance matrix $\mathbf{K}(\mathbf{x})$ and zero mean. The elements of $\mathbf{K}$ are determined by a covariance kernel function.

(a) Covariance function (1 point)

Implement a Matlab function `gp_cov(x1, x2, theta0, theta1, theta2, theta3, beta)`, that returns a $n1 \times n2$ covariance matrix. Use the following kernel function:

$$K(\mathbf{x}_n, \mathbf{x}_m) = \theta_0 \exp\left\{-\frac{\theta_1}{2}||\mathbf{x}_n - \mathbf{x}_m||^2\right\} + \theta_2 + \theta_3 \mathbf{x}_n^T \mathbf{x}_m + \beta^{-1}\delta_{nm}$$

(The last term is the nugget. See the lecture for an explanation.)

(b) Unconditioned Gaussian process (4 points)

Draw ten samples from a Gaussian process in the range `-1:0.1:1` and plot them as line plots using different colors (in a single figure). Use the covariance function from above and make three plots with the following three different parameter settings ($\theta_0, \theta_1, \theta_2, \theta_3, \beta$): `(9,4,0,0,0), (1,4,10,0,0), (0,0,0,5,0)`
Do more experiments with different parameter combinations on your own. Explain the influence of each of the four θ parameters on the resulting function.

(c) Conditioned Gaussian process (5 points)

We introduce three observed values at $x = (-0.1, 0.2, 0.8)$ and $y = (0.45, 1.0, -1.0)$. Draw 25 samples for the range `-2:0.05:2` from a Gaussian Process conditioned on these observed values. Use `(1,10,0,0,0)` as parameters for the covariance function. Plot the result together with the observed values (in a single figure).

(d) Nugget term (2 points)

Next, we activate the nugget term. Repeat the experiment from the previous subsection but use the parameters `(1,10,0,0,20)`). Plot your result. What is the effect of the nugget term? Give a theoretical explanation.

Prob. 2: Gaussian Process Regression

We again use the three observed values at $x = (-0.1, 0.2, 0.8)$ and $y = (0.45, 1.0, -1.0)$.

(a) Apply Gaussian process regression (2 points)

Plot the prediction of GP regression for all locations in `[-2:2]`. Use the covariance function with parameters `(1,theta1,0,0,25)` corresponding to an RBF kernel with a nugget term. Add confidence limits at plus or minus sqrt(prediction variance) to your prediction and plot them, too. Make three plots for `p=10,20,100`. Comment on the resulting predictions. Explain which model you would use, or what prevents you from making an informed decision.

(b) Plot observation weight functions (3 points)

Plot the weight functions (effective interpolation kernel) λ from $\hat{y} = Y * \lambda$ of each observation, in the range [-2, 2] for the RBF kernel (parameters `(1,10,0,0,25)`).

(c) Plot kernel weight functions (3 points)

For the same interval and the same kernel, plot the weighted RBF functions which, when summed up, give your estimate. That is, plot (YK^{-1}) from the equation $\hat{y} = YK^{-1}k$.

Regulations

Please hand in the matlab code, figures and explanations (describing clearly which belongs to which). Non-trivial sections of your code should be explained with short comments, and variables should have self-explanatory names. Plots should have informative axis labels, legends and captions. Please enclose everything into a single PDF document (e.g. use the publish command of MATLAB for creating a LaTeX document and run latex, dvips and ps2pdf or copy and paste everything into an office document and convert to PDF). Please email the PDF to patternrecognition@hci.iwr.uni-heidelberg.de before the deadline specified below. You may hand in the exercises in teams of two people, which must be clearly named on the solution sheet (one email is sufficient). Discussions between different teams about the exercises are encouraged, but the code must not be copied verbatim (the same holds for any implementations which may be available on the WWW). Please respect particularly this rule, otherwise we cannot give you a passing grade.