CPSC-540 Machine Learning 2005

Homework # 2

Due Thursday March 23 in class.

NAME:____

Signature:______________________________________

STD. NUM: ____________________________________

General guidelines for homeworks:

You are encouraged to discuss the problems with others in the class, but all write-ups are to be done on your own.
Homework grades will be based not only on getting the "correct answer," but also on good writing style and clear presentation of your solution. It is your responsibility to make sure that the graders can easily follow your line of reasoning.
Try every problem. Even if you can't solve the problem, you will receive partial credit for explaining why you got stuck on a promising line of attack. More importantly, you will get valuable feedback that will help you learn the material.
Please acknowledge the people with whom you discussed the problems and what sources you used to help you solve the problem (e.g. books from the library). This won't affect your grade but is important as academic honesty.
When dealing with Matlab exercises, please attach a printout with all your code and show your results clearly.

1. **Bayesian Linear Gaussian Inference**: If the posterior is proportional to the likelihood and prior:

$$p(\theta|\mathbf{X}, \mathbf{y}) \propto \left(\mathbf{2}\pi\sigma^{\mathbf{2}}\right)^{-\frac{\mathbf{n}}{\mathbf{2}}} \mathbf{e}^{-\frac{\mathbf{1}}{\mathbf{2}\sigma^{\mathbf{2}}}(\mathbf{y}-\mathbf{X}\theta)^{\mathbf{T}}(\mathbf{y}-\mathbf{X}\theta)} |\mathbf{2}\pi\sigma^{\mathbf{2}}\mathbf{R}^{-\mathbf{1}}|^{-\frac{\mathbf{1}}{\mathbf{2}}} \mathbf{e}^{-\frac{\mathbf{1}}{\mathbf{2}\sigma^{\mathbf{2}}}(\theta-\theta_{\mathbf{0}})^{\mathbf{T}}\mathbf{R}(\theta-\theta_{\mathbf{0}})},$$

rearrange terms in the exponents in order to obtain a simple analytical expression for the posterior distribution. In particular, show that it has the following **sufficient statistics**:

$$\begin{aligned} \mathbb{E}(\theta|\mathbf{X}, \mathbf{y}) &= (\mathbf{X}^{\mathbf{T}}\mathbf{X} + \mathbf{R})^{-\mathbf{1}}(\mathbf{X}^{\mathbf{T}}\mathbf{y} + \mathbf{R}\theta_{\mathbf{0}}) \\ cov(\theta|\mathbf{X}, \mathbf{y}) &= (\mathbf{X}^{\mathbf{T}}\mathbf{X} + \mathbf{R})^{-\mathbf{1}}\sigma^{\mathbf{2}} \end{aligned}$$

2. **Gaussian Processes**: Using the properties of the multivariate Gaussian distribution derived in Jordan's chapter, prove the expressions for the mean and variance of the predictions of a Gaussian process, when the mean function is known.

3. **Active Learning with GPs**: The course homework website has matlab functions demonstrating (i) A and D designs for the linear model, (ii) GPs for regression and (ii) active learning using GPs and graph kernels. Your task is to modify the GP code for regression to demonstrate A and D designs in the nonlinear setting. Generate the plots as is done in the linear model demo.