Task 1.2. Supervised Learning: Bayesian Linear Regressor

Garoe Dorta-Perez
CM50246: Machine Learning and AI

December 10, 2014

1 Introduction

Simple linear regressors are overconfident in their predictions. Bayesian linear regressors are an extension of this model used to solve this particular issue.

2 The problem

In this model we want to compute a posterior distribution given a set of training samples, as shown in Equation 1. Where $\mathbf{w}$ is a one dimensional array with the world state, $\mathbf{X}$ is a matrix with the data points, ϕ are the parameters of a linear function of the data.

$$Pr(\phi|\mathbf{X},\mathbf{w}) = \frac{Pr(\mathbf{w}|\mathbf{X},\phi)Pr(\phi)}{Pr(\mathbf{w}|\mathbf{X})} \tag{1}$$

The prior in Equation 1 is a normal distribution with 0 mean and spherical covariance. And the likelihood is a multivariate normal distribution, as shown in Equation 3. Where σ^2 is the covariance, $\mathbf{I}$ is the identity matrix and $\boldsymbol{\theta} = \{\phi, \sigma^2\}$.

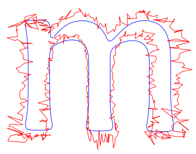
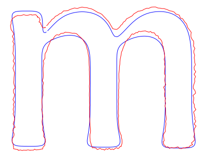
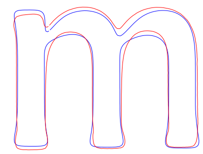
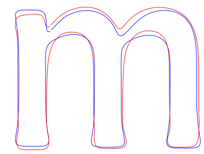
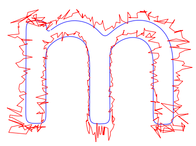
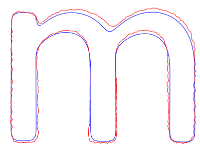
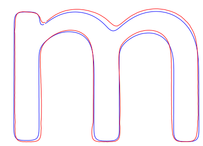
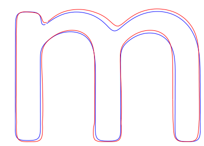
$$Pr(\mathbf{w}|\mathbf{X},\boldsymbol{\theta}) = Norm_{\mathbf{w}}\left[\mathbf{X}^T\phi, \sigma^2\mathbf{I}\right], \tag{2}$$

The posterior distribution is:

$$Pr(\phi|\mathbf{X},\mathbf{w}) = Norm_{\phi}\left[\frac{1}{\sigma^2}\mathbf{A}^{-1}\mathbf{X}\mathbf{w}, \mathbf{A}^{-1}\right], \tag{3}$$

$$\mathbf{A} = \frac{1}{\sigma^2}\mathbf{X}\mathbf{X}^T + \frac{1}{\sigma_p^2}\mathbf{I}, \tag{4}$$

3 Results

	$\sigma^2 = 10$	$\sigma^2 = 1$	$\sigma^2 = 0.01$	$\sigma^2 = 0.001$
Font1				
Font2				
Font3