Helsinki University of Technology
Lab of Computer and Information Science

Tik-61.184 Special Course in Information Technology
autumn 2000

Exercises, Session 8

1. Who was the Reverend Thomas Bayes?
2. Write a brief note on the historical controversy surrounding Bayesian statistics.
3. We require to fit the following data $\mathcal{D} = (\mathbf{x}^n, t^n)$ where t^n signifies the n^{th} target from a sample of N points, using a generalised linear model of the form $f(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T \phi(\mathbf{x})$. Using the Bayesian paradigm for learning we wish to infer the model parameters $\mathbf{w}$
   (a) Using Bayes rule give the expression for the posterior probability of the model parameters given the data and our belief in the appropriatness of using a GLM and associated basis function (denoted by $\mathcal{M}$).
   (b) We have a prior belief that the underlying function generating the data is smooth. Embody this belief using a prior distribution (with parameter α) on the $\mathbf{w}$ parameters.
   (c) Based on historical evidence we will assume that the errors in the model are normally distributed i.e. $\epsilon = t - \mathbf{w}^T \phi(\mathbf{x}) \sim \mathcal{N}_\epsilon(0, \beta)$ hence derive an expression for the posterior distribution of the parameters. Hint : Gaussian integrals will come in useful if your prior is conjugate to the error distribution.
   (d) Show that the mean of the posterior parameter distribution is
   $$E\{\mathbf{w}\} = \beta(\beta \Phi^T \Phi + \alpha \mathbf{I})^{-1} \Phi^T \mathbf{t} \tag{1}$$
   where $\Phi = [\phi(\mathbf{x}^1), \cdots, \phi(\mathbf{x}^N)]$ and $\mathbf{t} = [t^1, \cdots, t^N]$.
4. Discuss some of the problems which may be encountered when applying the same bayesian analysis to a nonlinear model of the form $f(\mathbf{x}, \mathbf{w}) = \varphi(\mathbf{w}^T \phi(\mathbf{x}))$ and how might these be overcome.