Homework 5: Linear Regression revisited

Solutions to this exercise sheet are to be handed in before the lecture on Friday, 23.11.12. Code (.m-file) and output (.mat-file, figures as .pdf-files) for the matlab-questions should be either submitted by email to `nicolas.ludolph@student.uni-tuebingen.de` (subject: [ML1] Exercise 5) or uploaded to Ilias. Use comments to explain your code. Please adhere to the file naming convention: `Homework5_<YourName>.<ext>`.

This exercise sheet will concentrate on linear regression and Gaussian random variables

1. (10 points) **Linear Regression (based on Bishop exercise 3.3)**. Consider a data-set in which each data point t_n has a weighting $r_n > 0$, so that the sum-of-square error function is

$$E(\omega) = \sum_{n=1}^{N} r_n (t_n - \omega^\top x_n)^2. \tag{1}$$

   (a) Find the parameter-vector $\hat{\omega}$ which minimizes this error function.

   (b) Describe two interpretation of this error functions in terms of i) replicated measurements and ii) a data-dependent noise-variance.

2. (20 points) **Regression with basis functions [matlab]** Download the file `Homework5.mat`, in which you will find training data `xTrain` (a vector of length $N = 20$) with outputs `tTrain`. Your job will be to train a nonlinear regression model from x to t using basis functions.

   (a) We want to use a 50-dimensional basis-set, i.e. the 'feature-vector' $z(x)$ should be 50-dimensional with $z_i(x) = 2\exp(-(x-i)^2/\sigma^2)$ with $\sigma = 5$ and $i = 1, \ldots 50$. Make a plot of the 50 basis functions (use the x-values in `xPlot`). Calculate the $50 \times N$ matrix `zTrain` for which the n-th row is $z(x_n)$, and produce an image of the matrix (using `imagesc`).

   (b) Using $\alpha = \beta = 1$ (same notation as in lectures), calculate the posterior mean $\mu = E(\omega|D)$ (a 50×1 vector) and plot it.

   (c) The posterior mean μ is a vector of weights of the basis functions. Calculate the corresponding predictive mean by $f_\mu(x) = E(t(x)|D) = \sum_{i=1}^{50} \mu_i z_i(x)$ and plot the predictive mean and the observed training data into the same plot.

   (d) Calculate the posterior covariance over weights $\Sigma = \mathrm{Cov}(\omega|D)$ and display it as an image. Extract the diagonal of Σ go obtain the posterior variance, and use it to plot $\pm$ 2 standard deviation error bars on the mean in part b)

   (e) [optional but recommended] Calculate, for each x (use the values in `xPlot`), the predictive variance $\mathrm{Var}(t|D, x)$, and use it to plot 'error bars' for the predictive distribution, i.e. $f_\mu(x) \pm 2\sqrt{\mathrm{Var}(t|D, x)}$.