Solution to Homework 2
Bayesian Methods

Anna Pavlova

Problem 1

To find a minumum of the error-function

$$E = \frac{1}{2} \sum_n r_n \left(t_n - \boldsymbol{\theta}^T \boldsymbol{x}_n\right)^2$$

(where $r_n > 0, \forall n$), set the gradient to zero

$$\nabla_\theta E = \sum_n r_n \left(\boldsymbol{\theta}^T \boldsymbol{x}_n - t_n\right) \boldsymbol{x}_n = 0$$

now since $\boldsymbol{a}^T \boldsymbol{b} = \boldsymbol{b}^T \boldsymbol{a}$, define a matrix $(\boldsymbol{X})_{ni} = (\boldsymbol{x}_n^T)_i$ and $\boldsymbol{R} = \mathrm{diag}(r_1, \ldots, r_N)$ have

$$\boldsymbol{X}^T \boldsymbol{R} \boldsymbol{X} \boldsymbol{\theta} - \boldsymbol{R} \boldsymbol{X}^T \boldsymbol{t} = 0$$

so the argmin is

$$\boldsymbol{\theta}^* = \left(\boldsymbol{X}^T \boldsymbol{R} \boldsymbol{X}\right)^{-1} \boldsymbol{R} \boldsymbol{X}^T \boldsymbol{t}.$$

This is truly a minimum, since all $r_n > 0$ and the functional is thus positive-definite (this is actually some weighted least-squares method). This weighting can be interpreted in two ways (i): different data points t_n may be affected by different level of noise (thus having different dispersions) and (ii): if some data points t_n appear several times, this can be simply tracked by considering one point weighted appropriately.

Problem 2

Need to proof inequality

$$\sigma_{N+1}^2(x) \leq \sigma_N^2(x)$$

We have variance $\sigma_N^2(x) = \frac{1}{\beta} + \phi(x)^T S_N \phi(x)$ and $S_N^{-1} = \alpha^{-1} I + \beta \Phi^T \Phi$ for calculation σ_{N+1}^2 found $S_{N+1}^{-}1$ by Woodbury matrix identity

$$\sigma_{N+1}^2(x):$$

$$\sigma_{N+1}^2(x) = \frac{1}{\beta} + \phi(x)^T S_{N+1} \phi(x)$$

$$S_{N+1}^{-1} = \alpha^{-1} I + \beta \Phi_{N+1}^T \Phi_{N+1}$$

$\Phi_{N+1}^T \Phi_{N+1} = \Phi_N^T \Phi_N + \phi_{N+1}\phi_{N+1}^T$ accordingly

$$S_{N+1}^{-1} = \alpha^{-1} I + \beta(\Phi_N^T \Phi_N + \phi_{N+1}\phi_{N+1}^T) = S_N^{-1} + \beta\phi_{N+1}\phi_{N+1}^T$$

found S_{N+1} using Woodbury matrix identity

$$(S_N^{-1} + \beta\phi_{N+1}\phi_{N+1}^T)^{-1} = S_N - \beta\frac{(S_N\phi_{N+1})(\phi_{N+1}^T S_N)}{1 + \beta\phi_{N+1}^T S_N \phi_{N+1}}$$

Let us prove that denominator ≥ 0, for that we have to note that $\phi^T S_N^{-1} \phi = \frac{||\phi||^2}{\alpha} + \beta||\Phi_N \phi||^2 \geq 0$. $S_N$is hermitian matrix, so $\phi^T (S_N \phi_{N+1})(\phi_{N+1}^T S_N)\phi = (\phi_{N+1}^T S_N \phi)^2 \geq 0$

$$\sigma_{N+1}(x) = \sigma_N(x) - \beta\frac{(\phi_{N+1}^T S_N \phi)^2}{1 + \beta\phi_{N+1}^T S_N \phi_{N+1}} \leq \sigma_N(x)$$