Homework 1

Due Date: March 22, 2018

Problem 1. Consider a data set in which each data point t_n is associated with a weighting factor $r_n > 0$, so that the sum-of-squares error function becomes

$$E_D(\mathbf{w}) = \frac{1}{2}\sum_{n=1}^{N} r_n\{t_n - \mathbf{w}^T\phi(\mathbf{x}_n)\}^2.$$

where $\phi(\mathbf{x}_n)$ is basis function. Find an expression for the solution $\mathbf{w}^*$ that minimizes this error function.

Problem 2. Assume we are given data$\{(\mathrm{x}_1,\mathrm{y}_1),...,(\mathrm{x}_n,\mathrm{y}_n)\}$ where x_i is a p-dimension vector and a parameter t>0. We denote by X the (n,p) matrix of row vectors x_1, x_2,...x_n and $\mathrm{y}=(\mathrm{y}_1,...,\mathrm{y}_n)^T$. The ridge regression estimator is defined as:

$$\widehat{\beta}_{ridge} = argmin_\beta \|y - X\beta\|^2 s.t. \|\beta\|^2 \leq t.$$

(a) Show that there exists a unique λ such that this formulation is equivalent to

$$\widehat{\beta}_{ridge} = argmin_\beta \|y - X\beta\|^2 + \lambda\|\beta\|^2$$

(b) Give the explicit form of the solution. Does it always exist?

Problem 3. Show that maximization of the class separation criterion given by:

$$m_2 - m_1 \propto \mathbf{w}^\mathbf{T}(\mathbf{m_2} - \mathbf{m_1})$$

with respect to w, using a Lagrange multiplier to enforce the constraint $\mathbf{w}^T\mathbf{w}= 1$, leads to the result that

$$\mathbf{w} \propto (\mathbf{m_2} - \mathbf{m_1})$$