Statistics 745

Group Assignment 8

1. Let $f(x_i) \in \{-1, 1\}$ for any x_i. Show that for any margin loss function $L(yf)$, the minimizer of:

$$\min_f \sum_{i=1}^{n} L(y_i f(x_i)) \tag{1}$$

   is solved by $y_i = f(x_i)$ for all i, i.e. interpolation.

2. Let $y_i \in \{-1, 1\}$. Denote W as the adjacency matrix of a graph. Define $\Delta = D - W$ where D is the row sum matrix of W. Answer the fallowing questions:

   (a) Show that Δ is positive semi-definite. Hint for any $\nu \neq \vec{0}$ show that $\nu^T \Delta \nu = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} (\nu_i - \nu_j)^2 W_{ij}}{2}$.

   (b) Let $L(y_i, f_i) = \log(1 + \exp(-2 y_i f_i))$. Derive an algorithm that solves the following:

$$\min_f \sum_{i=1}^{n} L(y_i, f_i) + f^T \Delta f$$

   (c) Implement your algorithm. This result is useful in that it can apply to observed graphs. Consider the cora text data (from the first group assignment). Let $W = A$ (the adjacency matrix for the network) and apply the algorithm to fit a function f over the text network. Compute the training accuracy with response as the papers topic.