582669 Supervised Machine Learning (Spring 2011)
Extra homework

This is completely voluntary set of problems that you may use to help in learning the topics that were covered after the last "real" homework set. Example solutions will be provided on the course web page, but you will not get any credit for these problems.

1. The *2-norm soft margin SVM* is like the 1-norm soft margin SVM except that the objective function is
$$-\mu + C\sum_{i=1}^{m} \xi_i^2$$
(the slack variables are squared). Show that now $\boldsymbol{\alpha}$ is obtained by maximising
$$W(\alpha) = -\frac{1}{4C}\sum_{i=1}^{m}\alpha_i^2 - \left(\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j k(x_i, x_j)\right)^{1/2}.$$

2. Assume that H_n is a class of binary classifiers with instance base $X_n = \{-1, 1\}^n$, for $n \in \mathbb{N}$. Show that $\log|H_n|$ is polynomial in n, if and only if $\mathrm{VCdim}(H_n)$ is polynomial in n.

   Hint: Use Sauer's Lemma (p. 163).

3. Let H_n be the concept class of n-dimensional axis-parallel rectangles ("boxes"). In other words, H_n consists of functions h such that for some constants $a_1, b_1, \ldots, a_n, b_n$ we have $h(x_1, \ldots, x_n) = 1$ if and only if $a_i \leq x_i \leq b_i$ for all i. Show that $\mathrm{VCdim}(H_n) = 2n$.

4. Let F be some class of functions from Z to $\mathbb{R}$. Given a sequence of functions $\boldsymbol{f} = (f_1, \ldots, f_n) \in F^n$ and a vector $\boldsymbol{v} \in \mathbb{R}^n$, define the function $\boldsymbol{v} \cdot \boldsymbol{f}$ by $(\boldsymbol{v} \cdot \boldsymbol{f})(z) = \sum_{i=1}^{n} v_i f_i(x)$. We can then define the *convex hull* of F as
$$\mathrm{conv}(F) = \left\{ \boldsymbol{v} \cdot \boldsymbol{f} \mid \boldsymbol{f} \in F^n, \boldsymbol{v} \in \mathbb{R}^n, \sum_{i=1}^{n} v_i = 1, v_i \geq 0 \text{ for all } i \right\}.$$
Show that $\mathrm{Rad}_m(F) = \mathrm{Rad}_m(\mathrm{conv}(F))$.

   Hint: Use Jensen on the absolute value function, and $\sup(f + g) \leq \sup f + \sup g$.