

GI01/4C55 – Homework #4 (Due 12am, December 5, 2005)

Aim: To get familiarity with boosting, support vector machines and kernels. Presentation, clarity, and synthesis of exposition will be taken into account in the assessment of these exercises.

This document is available at <http://www.cs.ucl.ac.uk/staff/M.Pontil/courses/SL-homework4.pdf>

1. **[40 pts] basic (boosting)** Write a computer program (preferably in Matlab or similar) which implements the boosting algorithm in the simplified scenario where the input data is two dimensional and the weak learners consist of decision stumps. Test the program on a simple dataset consisting of, say, 10 data points (you may choose a dataset similar to the one discussed in class). Report the training error of each weak learner, their weights, and the training error of the final classifiers.
2. **[30 pts] medium (SVM regression)** Consider the zero loss support vector machine regression problem:

$$\begin{aligned} \min_{\mathbf{w}, b} \quad & \mathbf{w}^\top \mathbf{w} \\ \text{subject to:} \quad & \mathbf{w}^\top \mathbf{x}_i + b - y_i \leq \epsilon \quad i = 1, \dots, m \\ & y_i - \mathbf{w}^\top \mathbf{x}_i - b \leq \epsilon \quad i = 1, \dots, m \end{aligned} \tag{1}$$

Derive the dual of this problem and discuss the property of the solution.

3. **[30 pts] advanced (reproducing kernel Hilbert spaces)**

Let $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ be a valid kernel function, that is a symmetric and positive semidefinite function (for example a Gaussian kernel). Consider the following space of functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ generated by K :

$$H_0 := \text{span} \left\{ K(\mathbf{x}, \cdot) : \mathbf{x} \in \mathbb{R}^d \right\} \equiv \left\{ f(\cdot) = \sum_{i=1}^{\ell} c_i K(\mathbf{x}_i, \cdot) : \ell \in \mathbb{N}, \mathbf{x}_i \in \mathbb{R}^d, c_i \in \mathbb{R}, i = 1, \dots, \ell \right\}.$$

This is clearly a linear space (if $f, g \in H_0$ and $a \in \mathbb{R}$ then $f + g \in H_0$ and $af \in H_0$). Consider the following inner product in H_0 : if $f(\cdot) = \sum_{i=1}^{\ell} c_i K(\mathbf{x}_i, \cdot)$ and $g(\cdot) = \sum_{i=1}^q d_i K(\mathbf{t}_i, \cdot)$ then

$$\langle f, g \rangle := \sum_{i=1}^{\ell} \sum_{j=1}^q c_i d_j K(\mathbf{x}_i, \mathbf{t}_j).$$

Show that this is indeed an inner product, that is it satisfies the following properties:

- (a) $\langle f, g \rangle = \langle g, f \rangle$
- (b) $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle$
- (c) $\langle af, g \rangle = a \langle f, g \rangle$
- (d) $\langle f, f \rangle \geq 0$
- (e) $\langle f, f \rangle = 0$ if and only if $f = 0$

(The first four properties are easy to verify. The last one requires more effort). Furthermore argue that the inner product is well defined, that is it does not depend on the parameterization of the functions f and g (by this we mean that even if, say, f can be written both as $\sum_{i=1}^{\ell} c_i K(\mathbf{x}_i, \cdot)$ and $\sum_{i=1}^{\ell'} c'_i K(\mathbf{x}'_i, \cdot)$ the inner product between f and some other function g does not depend on the specific parameterization chosen)

4. **[5pts BONUS]**

In the last question, note that $\langle K(\mathbf{x}, \cdot), K(\mathbf{t}, \cdot) \rangle = K(\mathbf{x}, \mathbf{t})$. Does this suggest a natural feature map for the kernel K ?