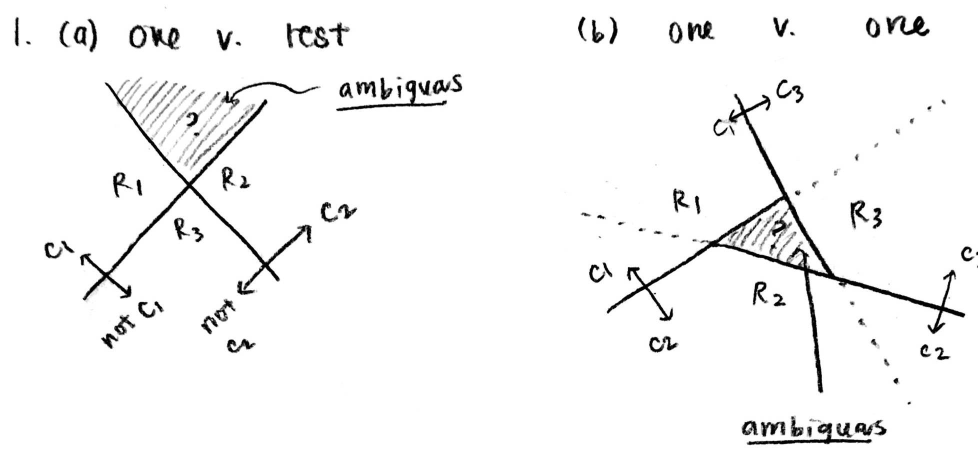
COMS W4771 Spring 2017: Homework #4

Kristy Choi

March 27, 2017

Question 1 Solutions

(a,b) The following illustration is an example of ambiguous regions that result from one vs. rest and one vs. one multi-class SVM:



where C_i denotes the "positive" class i and R_i denotes a particular classification region.

(c) The given specifications mirror the setup of (Crammer & Singer 2001). In developing the primal and dual of this multiclass SVM, we introduce the following notation:

- The set of m training data $S = \{(x_1, y_1), \ldots, (x_m, y_m)\}$
- The dimension of each x_i is n
- Each label y_i is an integer from the set $\mathcal{Y} = \{1, \ldots, k\}$
- $\boldsymbol{M} \in \mathbb{R}^{k \times n}$, where M_j is the jth row of $\boldsymbol{M}$ for $j = 1, \ldots, k$
- Slack variable ξ_i per data point, where $i = 1, \ldots, m$
- $\delta_{p,q}$, where $\delta_{p,q} = 1$ if $p = q$ and 0 otherwise.

Therefore the **primal (and quadratic) problem** is as follows:

$$\min_{M,\xi} \frac{1}{2}||M||_2^2 + \sum_{i=1}^{m} \xi_i$$
$$\text{subject to: } \forall i, j,\ M_{y_i} \cdot x_i + \delta_{y_i,j} - M_j \cdot x_i \geq 1 - \xi_i$$

We use the KKT conditions and Lagrange multipliers to solve for the dual (refer to p. 270 - 272 for the math). Some notation:

- The kernel function $K(.,.)$
- $\bar{\mathbf{1}}$ is the vector whose components are all 1
- $\mathbb{1}_i$ is the zero vector whose ith component is 1
- $\bar{\eta}_i$ = the $k-$dimensional vector corresponding to the ith data point in the dual variable η (one dual variable $\eta_{i,j}$ is added for each constraint)
- $\tau_i = \mathbb{1}_{y_i} - \bar{\eta}_i$

Therefore the **dual problem** is as follows:

$$\max_{\tau} -\frac{1}{2}K(x_i, x_j)(\tau_i, \tau_j) + \sum_i \tau_i \cdot \mathbb{1}_{y_i}$$
$$\text{subject to: } \forall i, \tau_i \leq \mathbb{1}_{y_i} \text{ and } \tau_i \cdot \bar{\mathbf{1}} = 0$$