

Homework 3: Support Vector Machines

Lecturer: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

Exercise 1. (★★) Consider a training data set $\mathcal{D} = \{z_i = (x_i, y_i)\}_{i=1}^m$. Consider the Soft-SVM Algorithm that requires the solution of the following quadratic minimization problem (in a slightly modified but equivalent form to what we have discussed)

Primal problem:

$$(0.1) \quad (w^*, b^*, \xi^*) = \arg \min_{(w, b, \xi)} \left(\frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^m \xi_i \right)$$

$$(0.2) \quad \text{subject to: } y_i (\langle w, x_i \rangle + b) \geq 1 - \xi_i, \quad \forall i = 1, \dots, m$$

$$(0.3) \quad \xi_i \geq 0, \quad \forall i = 1, \dots, m$$

for some user-specified fixed parameter $C > 0$.

- (1) Specify the Lagrangian function L associated to the above primal quadratic minimization problem, where $\{\alpha_i\}$ are the Lagrange coefficients wrt (0.2), and $\{\beta_i\}$ are the Lagrange coefficients wrt (0.3). Write down any possible restrictions on the Lagrange coefficients.
- (2) Compute the dual Lagrangian function denoted as \tilde{L} as a function of the Lagrange coefficients and the data points \mathcal{D} .
- (3) Apply the Karush–Kuhn–Tucker (KKT) conditions to the above problem, and write them down.
- (4) Derive and write down the dual Lagrangian quadratic maximization problem, along with the inequality and equality constraints, where you seek to find $\{\alpha_i\}$.
- (5) Justify why the i -th point x_i lies on the margin boundary when $\alpha_i \in (0, C)$ (beware it is $\alpha_i \neq C$), and why the i -th point x_i lies inside the margin when $\alpha_i = C$.
- (6) Given optimal values $\{\alpha_i^*\}$ for Lagrangian coefficients $\{\alpha_i\}$ as they are derived by solving the dual Lagrangian maximization problem in part 4, derive the optimal values w^* and b^* for the parameters w and b as function of the support vectors. Regarding parameter b it should be in the derived in the form

$$b^* = \frac{1}{|\mathcal{M}|} \sum_{i \in \mathcal{M}} \left(y_i - \sum_{j \in \mathcal{S}} \alpha_j^* y_j \langle x_j, x_i \rangle \right)$$

where you determine the sets \mathcal{M} and \mathcal{S} .

- (7) Report the halfspace predictive rule $h_{w,b}(x)$ of the above problem as a function of α^* and b^* .

Solution.