Machine Learning, Tutorial 5
University of Bern

Mehdi Noroozi (noroozi@inf.unibe.ch)

24/10/2018

Constraint Optimization and SVM

1. Consider the constrained minimisation problem below. Solve it using the KKT conditions.

$$\begin{aligned} \min_{x,y} \quad & x^2 + y^2 \\ \text{subject to} \quad & (x-3)^2 + y^2 \leq 4 \end{aligned}$$

2. Let $\{x^{(i)}, y^{(i)}\}_{i=1}^m$ be a set of m training examples of feature vectors $x^{(i)}$ and corresponding labels $y^{(i)}$. We consider binary classification, and assume $y^{(i)} \in \{-1, +1\}$ for $i = 1, \ldots, m$. The following is the primal formulation of L^2-SVM, a variant of the standard Support Vector Machine (SVM) formulation obtained by squaring the hinge loss:

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^m \xi_i^2 \\ \text{s.t.} \quad & y^{(i)}(w^\top x^{(i)} + b) \geq 1 - \xi_i,\ i = 1, \ldots, m, \\ & \xi_i \geq 0,\ i = 1, \ldots, m, \end{aligned}$$

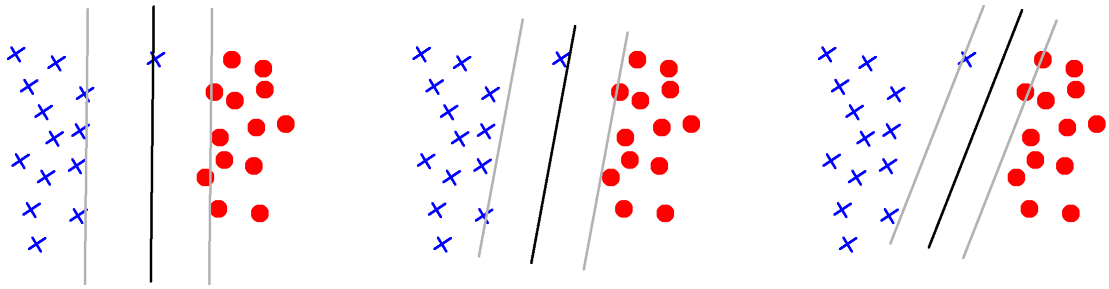
where $w,\ b$ are parameters of the usual SVM binary classifier model, ξ are slack variables, and $C > 0$ is a tuning parameter.

(a) Show that removing the last set of inequality constraints $\{\xi_i \geq 0\}_{i=1}^m$ does not change the optimal solution of the above primal formulation.
Hint: Consider how the constraints $y^{(i)}(w^\top x^{(i)} + b) \geq 1 - \xi_i$ change when ξ_i is positive or negative.

(b) After removing the last set of inequality constraints we arrive at the simplified problem:

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^m \xi_i^2 \\ \text{s.t.} \quad & y^{(i)}(w^\top x^{(i)} + b) \geq 1 - \xi_i,\ i = 1, \ldots, m. \end{aligned}$$

Write the dual formulation of this problem.
Hint: Recall that the dual formulation involves the Lagrangian of the primal problem, and that the Lagrangian is some linear combination of the objective function and of the constraints.



a) b) c)

3. The regularized Support Vector Machine minimizes the following objective,

$$\begin{aligned} \min_{\xi,w,b} \quad & \tfrac{1}{2}||w||^2 + C\sum_{i=1}^{m}\xi_i \\ \text{s.t.} \quad & y^{(i)}(w^T x^{(i)} + b) \geq 1 - \xi_i, \quad i = 1, ..., m \\ & \xi_i \geq 0, \quad i = 1, ..., m. \end{aligned}$$

The figures above show different solutions $a)$, $b)$ and $c)$ obtained with different C parameters.

Which solution belongs to $C = 0.01$, $C = 1$ and to $C = 100$? Briefly justify your answer.

4. The following image shows some training data $x_i \in R^2$, and $y_i \in \{-1, +1\}$. Circles represent the positive, and crosses represent the negative examples.

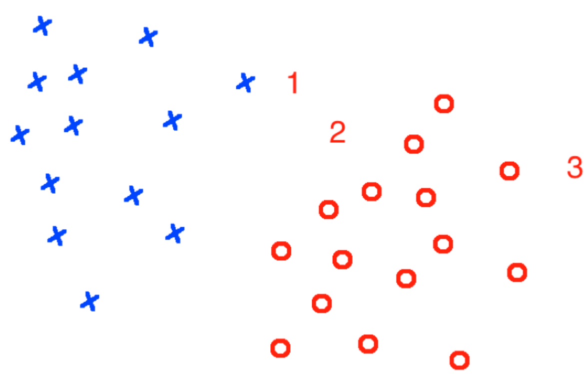


Figure 1: Train data.

- Which training points are likely to be support vectors?

- We add another positive training example to the training set. What happens to w and the margins, if we place the training example at location 1,2 or 3?