Note: Support Vector Machines

Sun Zhao

January 1, 2013

1 Support Vector Machines Intuition

A support vector machine(SVM for short) constructs a hyperplane or set of hyper-planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier. H_1, H_2, H_3 in Fig. 1 are three hypothesises trying to separate black and white pointers. Obviously, H_1 does not separate the classes. H_2 does, but only with a small margin. H_3 separates them with the maximum margin. SVM chooses the hyperplane so that the distance from it to the nearest data point on each side is maximized. Recall that in logistic regression, we use $\Theta^T x = 0$ as the hyper-plane, and pre-

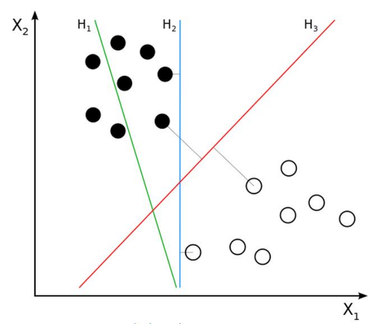

Figure 1:

dict positive class if $\Theta^T x \geq 0$, otherwise, negative class. SVM is trying to keeps a minimum margin between the two classes and wants $\Theta^T x \geq 1$ if x is positive and $\Theta^T x \leq -1$ if x is negative. Let H_1 denotes $\Theta^T x = 1$, H_0 denotes $\Theta^T x = -1$ and H_0 denotes $\Theta^T x = 0$. The margin distance between H_1 and H_{-1} is $\frac{2}{||\Theta_{1\ldots n}||}$ where $||\Theta_{1\ldots n}||$ equals $\sqrt{\sum_{i=1}^{n} \Theta_i^2}$. The location relationships between training examples and H_1, H_0, H_{-1} is shown in Fig. 2. The optimization problem can be summarized as follow:

$$\min ||\Theta_{1\ldots n}|| \tag{1}$$

Subject to(for any i = 1 ... m)

$$y^{(i)}(\Theta^T x^{(i)}) \geq 1 \tag{2}$$

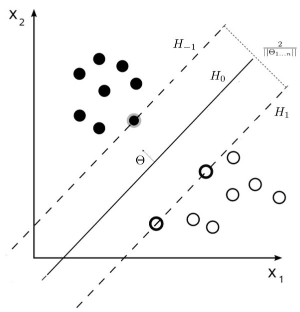

Figure 2:

2 Cost Function

Instead of following the logarithm example cost function, SVM introduces cost function show in Fig. 3. The new cost functions have following properties:

if $y^{(i)}$=1, cost = 0 when $\Theta^T x^{(i)} \geq 1$
if $y^{(i)}$=0, cost = 1 when $\Theta^T x^{(i)} \leq -1$

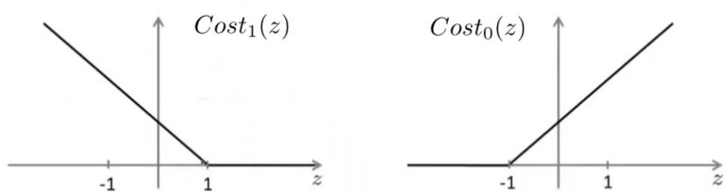

Figure 3: