

## 6.867 Problem Set 2

October 19, 2015

### Support Vector Machines

Let's now compare the performance of SVM to that of logistic regression for classification problems. To illustrate the objective and constraints of the support vector machine, we have included below the explicit objective we would optimize over, as well as the constraints, for the dual form of a linear SVM with slack variables. The equations below correspond to the 2D problem where we have positive examples (1, 2), (2, 2) and negative examples (0, 0), (-2, 3).

$$\min_{\alpha_1, \alpha_2, \alpha_3, \alpha_4} \frac{1}{2} [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4] \begin{bmatrix} 5 & 6 & 0 & -4 \\ 6 & 8 & 0 & -2 \\ 0 & 0 & 0 & 0 \\ -4 & -2 & 0 & 13 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} + [-1 \quad -1 \quad -1 \quad -1] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}$$

s.t.

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ C \\ C \\ C \\ C \end{bmatrix},$$

$$[1 \quad 1 \quad -1 \quad -1] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = 0$$

Table 1: Performance of Linear SVM on provided data sets

dataset	classification error rate (training set)	classification error rate (validation set)
data_stdev1	0.00%	0.00%
data_stdev2	9.50%	7.50%
data_stdev4	26.00%	23.50%
data_nonsep	49.50%	51.25%

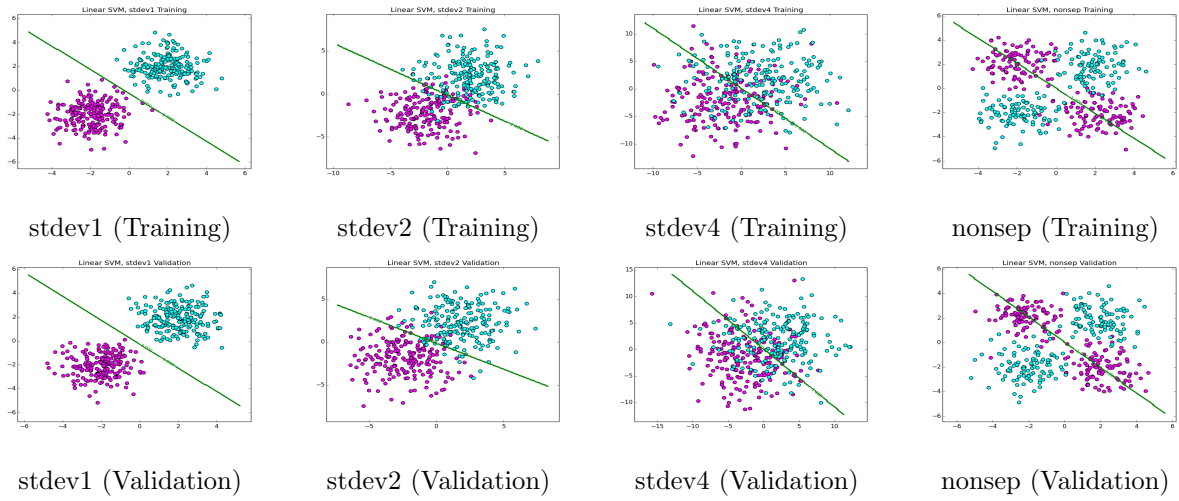


Figure 1: The decision boundaries generated by SVM plotted against the stdev1, stdev2, stdev4, and nonsep (training and validation) datasets