

ST790 Homework 6 Solution

1. (Ex 4.6)

- (a) Separability means there exists a β such that $y_i \beta^T z_i > 0, \forall i$. Let $c = \min_i y_i \beta^T z_i$, then $\beta_{\text{sep}} = \beta/c$ satisfies $y_i \beta_{\text{sep}}^T z_i > 0, \forall i$.

2. (Two-Class Classification Problem: Scenario 2)

For this and the next question, we use the `e1071::svm` function to fit the training data. The `cost` parameter, aka γ in the slides, is tuned and other parameters are prespecified. Specifically, the Gaussian kernel is $K(\mathbf{x}, \mathbf{x}') = \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / \dim(\mathbf{x}))$ and the polynomial kernel is $K(\mathbf{x}, \mathbf{x}') = \exp(1 + \langle \mathbf{x}, \mathbf{x}' \rangle)^3$, where $\dim(\mathbf{x})$ is the dimension of the input features, which is 2 in this problem.

Tuning all the parameters together is also possible using grid search but tuning one parameter at a time without several iterations will not yield the optimal combination.

The optimal γ and test errors for each SVM classifier are listed as follows (the answer may vary depending on the data generating process).

Kernel	γ	Test Error
Linear	0.048	0.243
Gaussian	0.056	0.225
Polynomial	0.091	0.233

3. (Two-Class Classification Problem: Zip Code Data)

The optimal γ and test errors for each SVM classifier are listed as follows.

Kernel	γ	Test Error
Linear	0.001	0.025
Gaussian	1.000	0.025
Polynomial	0.100	0.025