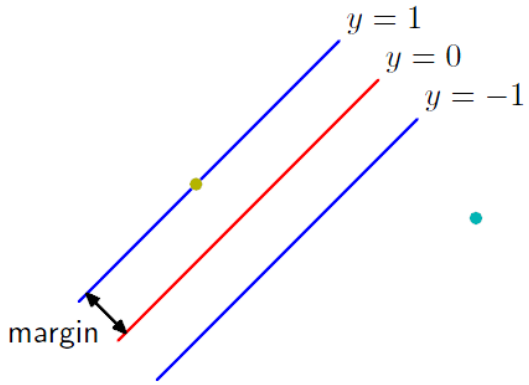
CS 229 ASSIGNMENT 7:

Support Vector Machine

NAME: Xiaopeng Xu KAUST ID: 129052

Question1: (30 points) Margin for the maximum-margin hyper-plane (Exercise 7.4 of Bishop's book, 0 pt for solutions copied from internet)



Show that the value ρ of the margin for the maximum-margin hyperplane is given by

$$\frac{1}{\rho^2} = \sum_{n=1}^{N} a_n$$

where $\{a_n\}$ are given by maximizing

$$\tilde{L}(\mathbf{a}) = \sum_{n=1}^{N} a_n - \frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N} a_n\, a_m t_n t_m k(\mathrm{x}_n, \mathrm{x}_m)$$

subject to

$$a_n \geq 0, \qquad n = 1 \ldots N$$

and

$$\sum_{n=1}^{N} a_n\, t_n = 0$$

Answer is in attachment 1.

Question2: (70 points) Classification by SVM

Code:

Use the LIBSVM http://www.csie.ntu.edu.tw/~cjlin/libsvm/ Support Vector Machine

library for the classification.

Data:

You can download the data from http://archive.ics.uci.edu/ml/datasets/Wine+Quality
Take either the red-wine or the white-wine data set.
Take "quality" as class label, e.g., 1-5 as **negative**, while 6-10 as **positive**.

Evaluation and Testing:

Divide the whole data set into training data and testing data, e.g., 60% for training and 40% for testing.
Use 5-fold Cross Validation for setting parameters, e.g., C and kernel parameters.

Questions to answer:

We have 3 different types of models for learning classifiers (SVM with 3 different types of kernel):

1 SVM with linear kernel
2 SVM with polynomial kernel
3 SVM with the radial basis function kernel

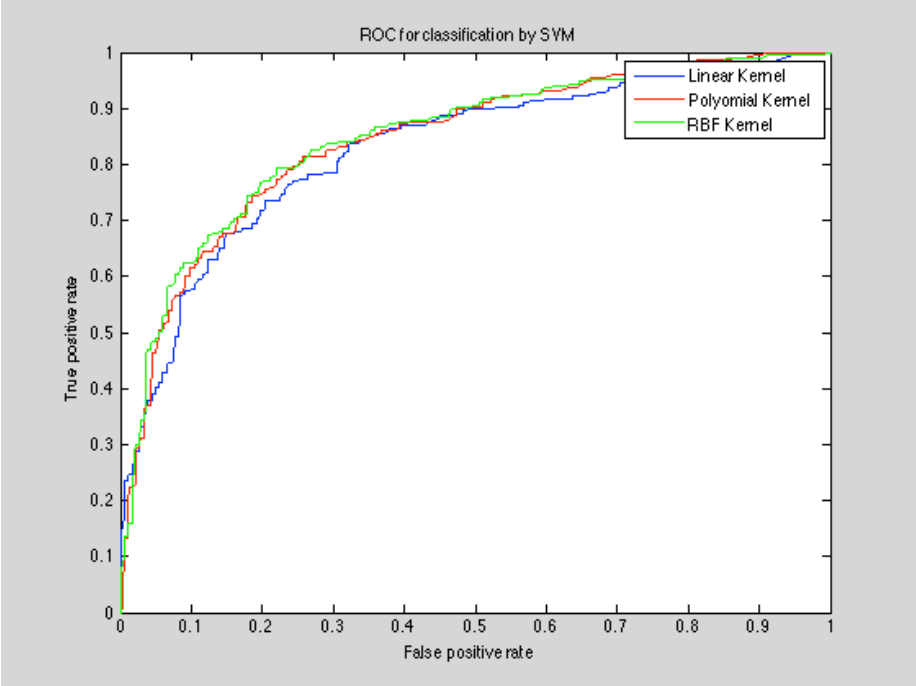
(1) (15 points) For each learning model (each type of kernel), use 60% of data for training SVM model (with the default parameters), and use the remaining 40% for testing.

a) Report the number of support vectors
b) Plot the ROC curve of testing results by ranking the decision values (3 curves in one figure)
c) Compute the AUC (Area Under Curve), which kernel is the best?

I ran each kernel with default params for 5 times, and then take mean to evaluate them.

Mean of 5 trials	Linear Kernel	Polynomial Kernel	RBF Kernel
No. of Support Vectors	566	658	593
Accuracy	73.3906	75.0168	75.4546
AUC	0.8115±0.0104	0.8236±0.0158	0.8317±0.0145

ROC Curve:



(2) (45 points) For each learning model (each type of kernel),

Use 5-fold cross validation for setting the parameters of training process. Please note different kernels may have different parameters to set.

After cross validation, choose the best parameter setting, train the model by 60% of data again (the same data used in (1)), test the model by the remaining 40% of data.

a) Report the setting of parameters
b) Report the number of support vectors
c) Plot the ROC curve of testing results by ranking the decision values (3 curves in one figure)
d) Compute the AUC (Area Under Curve), which kernel is the best?

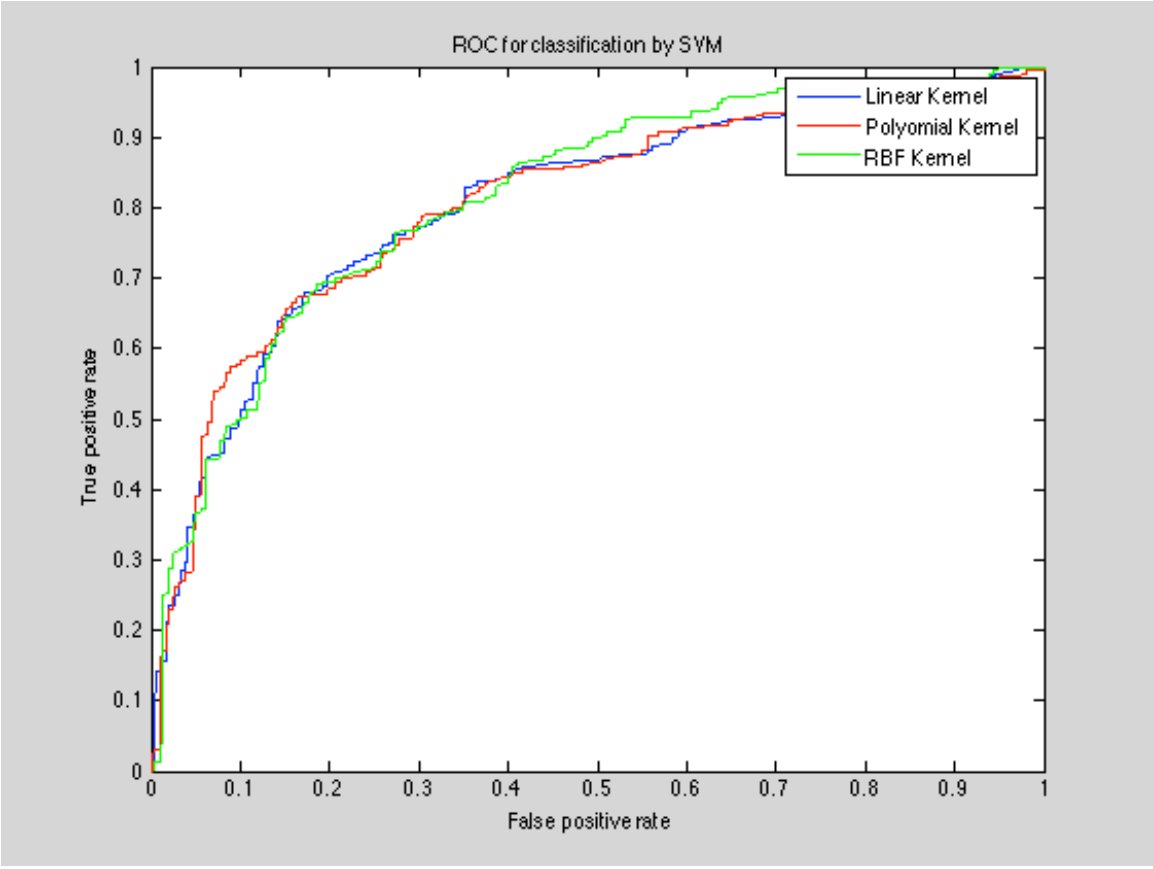
I tried each kernel with default parameters for 5 times, and then take mean to find best parameters. Details of the parameters I tried are in the code.

Parameters for linear kernel:`'-t 0 -c 1'`
Parameters for polynomial kernel: `'-t 1 -c 1 -d 3 -g 0.01 -r 10'`
Parameters for rbf kernel:`' '-t 2 -c 10 -g 1/11'`

Number of Support Vectors and AUC are taken from 5 runs.

Mean of 5 trials	Linear Kernel	Polynomial Kernel	RBF Kernel
No. of Support Vectors	567	548	819
AUC	0.8138±0.0172	0.8225±0.0155	0.8233±0.0138

ROC Curve:



(3) (10 points) make a table for comparing the AUC of different kernels with

different setting of parameters (totally 6 AUC values), report the 3 best models (decided by their AUC values).

Model	AUC
Linear (Default)	0.8115±0.0104
Polynomial (Default)	0.8236±0.0158
RBF (Default)	0.8317±0.0145
Linear (Select)	0.8138±0.0172
Polynomial (Select)	0.8225±0.0155
RBF (Select)	0.8233±0.0138

From this I find RBF kernel with default parameters, RBF kernel with selected parameters, and Polynomial kernel with default parameters gives me best AUC values.