a. **Using CVX package**

The optimization of SVM dual is implemented in function *p2*. By derivation and comparing to the original dual, the weight vector w and the intercept term b are given by:

$$Q_{ij} = -0.5 * Y_i * Y_j * \left(X_i^T * X_j \right)$$

And in matrix form, the Q matrix can be given by

$$Q = -0.5 * \left(Y * Y^T \right) .* \left(X * X^T \right)$$

Similarly, b and c were obtained to be

$$b = ones\left(m, 1 \right)$$

$$c = 0$$

The number of support vectors with C=1 obtained are displayed in the table shown below. The complete output of the function with the indices of the respective support vectors are present in the files 'a_small(Linear) .txt', 'a_small(Rbf) .txt', 'a_bigcvxLin.txt',' a_bigcvxRbf.txt' respectively.

Type of Kernel/File	Number of features	Number of support vectors
Linear Kernel (with train-small)	996	152
Gaussian Kernel (with train-small)	996	258
Linear Kernel (with train)	1000	452
Gaussian Kernel (with train)	1000	752

VARIATION OF NUMBER OF SUPPORT VECTORS

Observations:

b. **Weight vector w and the intercept term b**:

By calculating the weights and intercepts, the accuracies were obtained as follow:

Size of input	Intercept Term	Accuracy (%)
Linear Kernel (with train-small)	0.0832	91.3 % (913/1000)
Linear Kernel (with train)	-0.08522	98.7 % (987/1000) *
Gaussian Kernel (with train-small)	0.5047	87 % (870/1000)
Gaussian Kernel (with train)	0.3394	96.9 % (969/1000) *

ACCURACIES AND INTERCEPT TERMS WITH THE SIZE OF FILE IN LINEAR AND GAUSSIAN KERNELS

Observations:

- The above result justifies that with bigger training set the SVM for spam classification learnt much better as compared to train-small file.
- Linear kernel performs marginally better than the Gaussian kernel as clear from the accuracy values.

c. **Using the Gaussian kernel:**
The number of support vectors and the accuracies obtained have already been shown in the tables above. Now the Q matrix is changed such that

$$Q_{ij} = -0.5 * Y_i * Y_j * K(X_i, X_j)$$

where

$$K(X_i, X_j) = e^{-\gamma * \|X_i - X_j\|^2}$$

Whereas b and c are not changed.
The Gaussian and Linear kernel for the big file took almost 6 hr each on GCL machines for this part.

d. **Using the LibSvm Library:**
The outputs for this part are stored in files ‘d_gauss_libsvm_big.txt’, ‘d_gauss_libsvm_small.txt’, ‘d_lin_libsvm_small.txt’ and ‘d_linear_libsvm_big.txt’.

Size of input	Accuracy (%)	Number of support

		vectors
Linear Kernel (with train-small)	91.3 % (913/1000)	152
Linear Kernel (with train)	98.7 % (987/1000)	452
Gaussian Kernel (with train-small)	89.2% (892/1000)	530
Gaussian Kernel (with train)	98.7 % (987/1000)	1982

ACCURACIES AND SVS WITH DIFFERENT PARAMETERS USING LIBSVM

Observations:

- The LibSVM performs better as compared to CVX implementation of the same problem. This is because CVX is general purpose quadratic optimization software but LibSVM is highly optimized for support vector learning.
- The number of support vector using CVX and LibSVM were almost found to be same in number as given in plot. This cross verifies the CVX implementation also.
