183.605
Machine Learning for Visual Computing
Assignment 2

Group 12:
Hanna Huber (0925230)
Lena Trautmann (1526567)
Elisabeth Wetzer (0726681)

January 18, 2017

1 Assignment 2

1.1 The dual optimization problem

Tasks:

- Generate a suitable training set of linearly separable data

The training set is generated in `generateTrainingData.m`. The function `generateTrainingData(N, xRange, yRange, linear)` takes the number of sample points, the domains (defined by a lower and an upper bound) from which the x and y coordinates are sampled and a flag which indicates if the resulting data should be linearly separable. A set of random 2D coordinates is created using the MATLAB function `rand`. Linear separability is achieved by labelling the according to the condition $x_i + y_i > \bar{x} + \bar{y}$, where $\bar{x}$ and $\bar{y}$ denote the domain centers.

- Plot the input vectors in $\mathbb{R}^2$ and visualize corresponding target values (e.g. by using color).

The resulting training data is displayed in Figure 1. Sample points with class label 1 are marked red, points with label -1 are merked green.

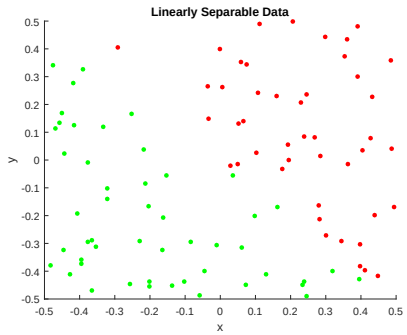


Figure 1: Linearly separable data with color-coded class labels (1=red, -1=green) for $N = 100$.

- Visualize the support vectors and plot the decision boundary.

Figure 2 shows the support vectors defined by `trainSVM`.

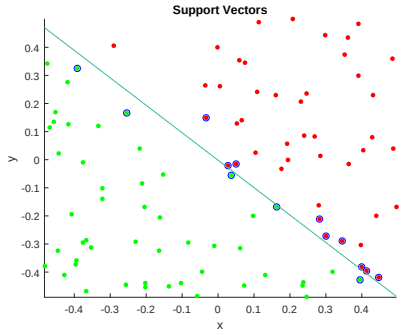
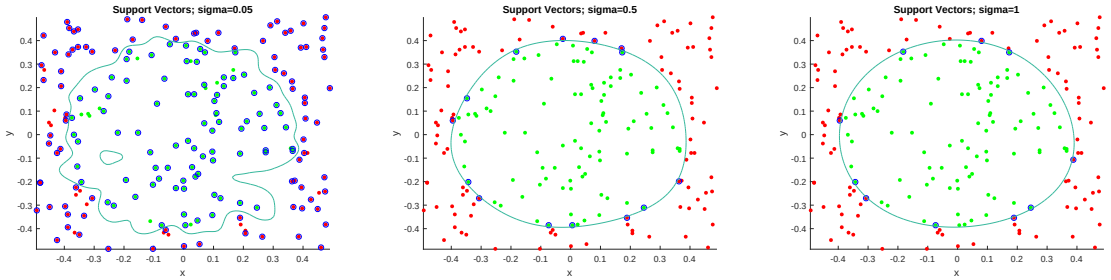


Figure 2: Training data with support vectors marked by blue circles and decision boundary plotted in blue.



(a) $\sigma = 0.05$ (b) $\sigma = 0.5$ (c) $\sigma = 1$

(d) $\sigma = 4$ (e) $\sigma = 7.06$ (f) $\sigma = 7.07$

Figure 3: Training data with support vectors marked by blue circles and decision boundary plotted in blue, for different values of the RBF parameter σ.

1.2 The kernel trick

Tasks:

- Try different values for σ (the RBF parameter).

The radial basis function kernel is defined by $K(x, y) = exp(-\frac{\|x-y\|^2}{\sigma^2})$. The corresponding support vectors are shown in Figure 3 for different values of the RBF parameter σ. We can see, that the decision boundary gets better for larger σ up to a value of 7.07. With $\sigma >= 7.07$ the values of α are all greater than 10^{-8}, hence we have all data points as support vectors. Furthermore, the least number of SVs was determined for $\sigma = 7.06$.

- Generate a non-linearly separable training set, plot the data, visualize the support vectors and plot the decision boundary.

See also Figure 3