HÁSKÓLINN Í REYKJAVÍK
REYKJAVIK UNIVERSITY

Engineering Department

Gagnanám og Vitvélar

T-809-DATA

Assignment 8

Bjarki Laxdal
Email: Bjarki18 AT ru.is
Phone: 843-9292

9th October 2022

1 Section 1.1

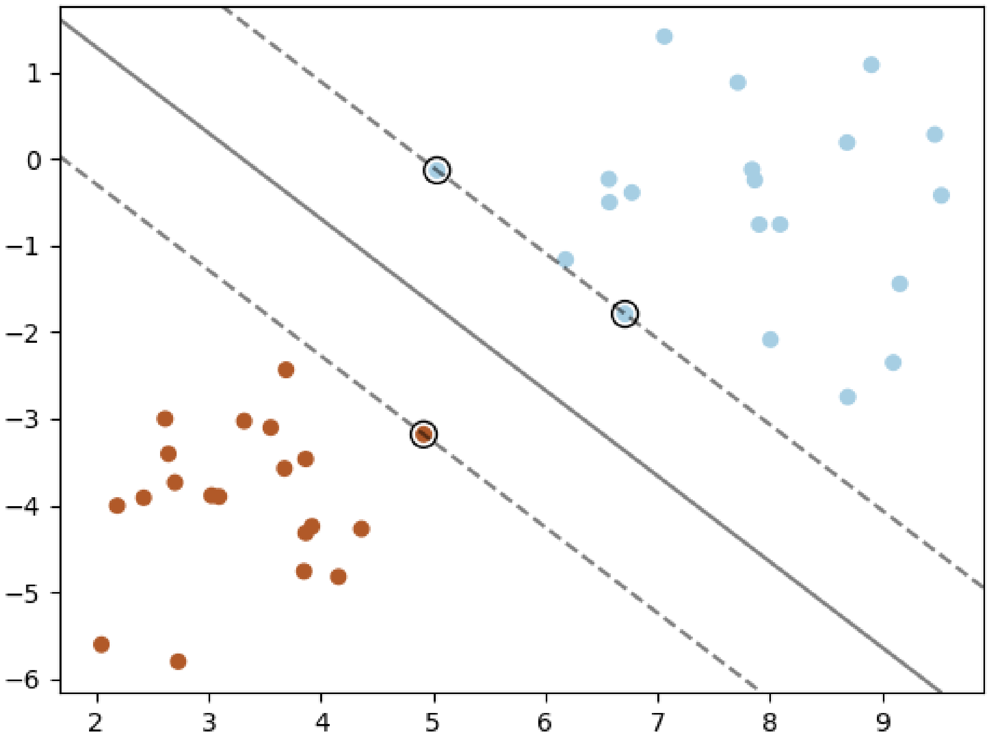

Figure 1: Linear kernel

2 Section 1.2

- How many support vectors are there for each class in your example?
  - According to the SVM algorithm we find the points closest to the line from both the classes. These points are called support vectors. In our case for the blue points there are 2 support vectors, and for the red points there is one support vector.
- What is the shape of the decision boundary?
  - The decision boundary is linear because the kernel is linear.

3 Section 1.3

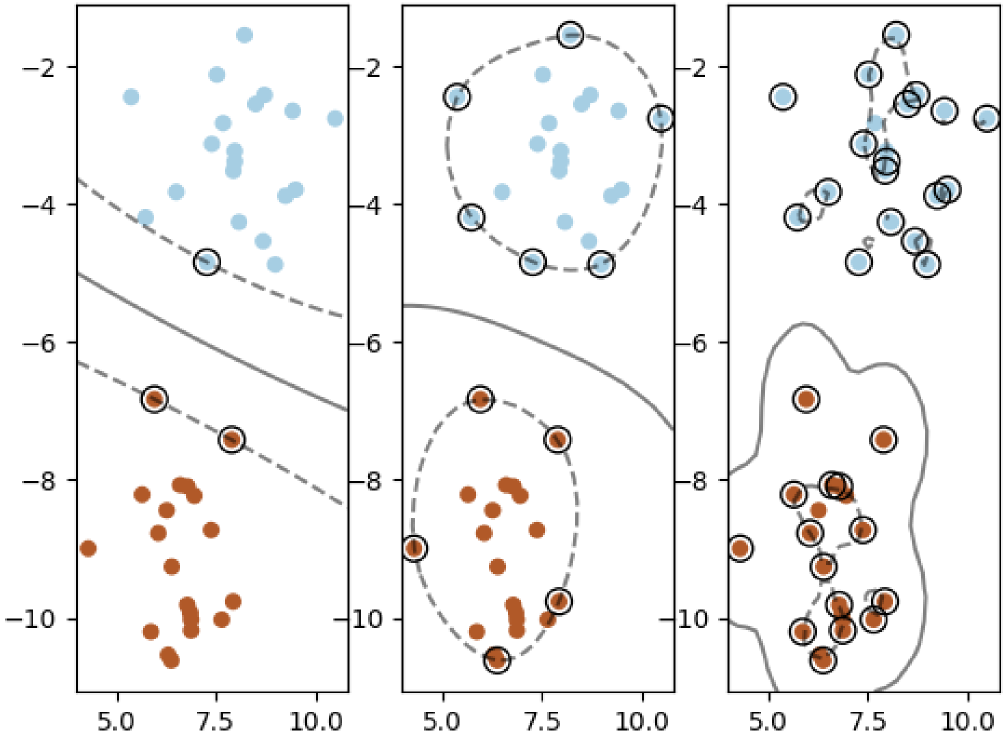

Figure 2: Radial basis function, from the left, gamma = default, gamma = 0.2, gamma = 2

4 Section 1.4

- How many support vectors are there for each class for each value of gamma?
  - Gamma = default -> 1 blue, 2 red
  - Gamma = 0.2 -> 6 blue, 5 red
  - Gamma = 2 -> 18 blue, 15 red.
- What is the shape of the decision boundary for each value of gamma?
  - The shape of the decision boundaries are all polynomial, where as the gamma increases the degree of the polynomial increases. The higher the value of gamma, the more influence the data points will have on the decision boundary. The decision boundary follows the shape of the hyperplane, which in an n-dimensional Euclidean space is a flat, n-1 dimensional subset of that space that divides the space into two disconnected parts.

- What difference does the gamma parameter make and why?
  - It defines how far the influence of data points reaches. If we have a high value of gamma, then the decision boundary will be dependant on the points that are close to the line, which means that we will ignore some points that are far away from the decision boundary. This is because the closer the points are, the more weight they are given and it will result in a squiggly decision boundary, much like the one for gamma = 2 in figure 2 (right-most). The opposite happens when the gamma value is low, in which the data points further away get a much larger weight, which results in a more linear curve, like the one for gamma = default, in figure 2(left-most).

5 Section 1.5

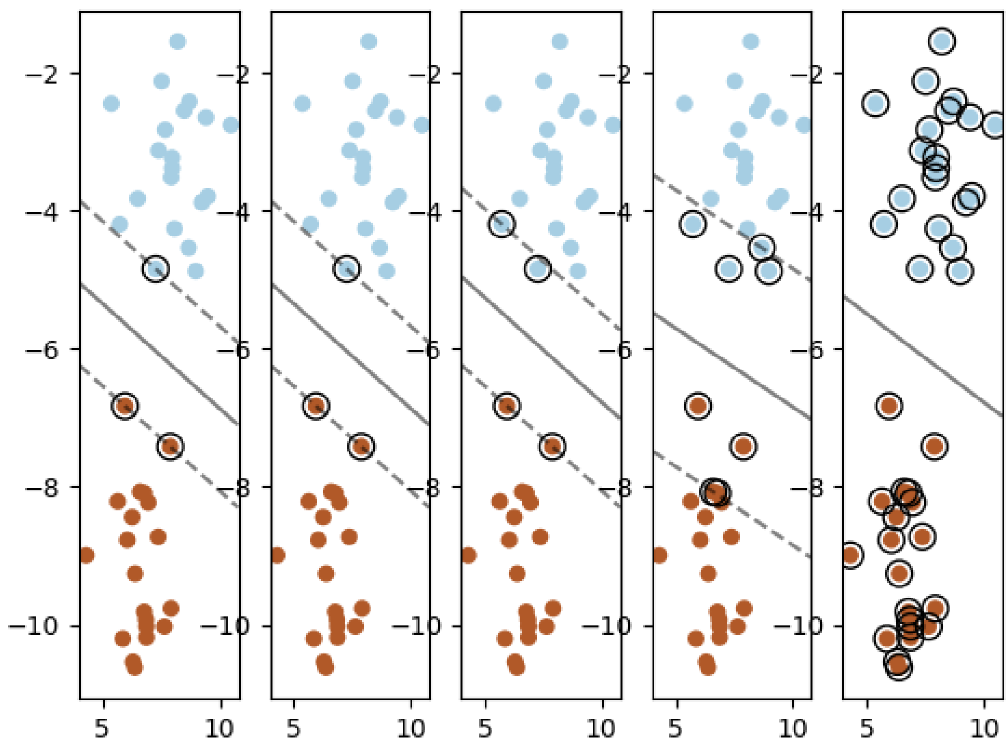

Figure 3: Different values of C, linear kernel, from left to right C = 1000, C = 0.5, C = 0.3, C = 0.05, C = 0.001.

6 Section 1.6

- How many support vectors are there for each class for each case of C?
  - C = 1000 -> 1 blue, 2 red
  - C = 0.2 -> 1 blue, 2 red
  - C = 0.3 -> 2 blue, 2 red
  - C = 0.05 -> 4 blue, 4 red
  - C = 0.001 -> 20 blue, 20 red
- How many of those support vectors are within the margins?.
  - C = 1000 -> 0 blue, 0 red
  - C = 0.2 -> 0 blue, 0 red
  - C = 0.3 -> 0 blue, 0 red
  - C = 0.05 -> 3 blue, 2 red
  - C = 0.001 -> 20 blue, 20 red
- Are any support vectors misclassified? If so, why?
  - None of the support vectors are misclassified, since no points were misclassified, this tells us that the classes are separable.

7 Section 2.2

- Linear
  - Accuracy = 0.9239766081871345
  - Precision = 0.9320388349514563
  - Recall = 0.9411764705882353
- Radial
  - Accuracy = 0.9122807017543859
  - Precision = 0.9223300970873787
  - Recall = 0.9313725490196079
- Polynomial
  - Accuracy = 0.9005847953216374
  - Precision = 0.9047619047619048
  - Recall = 0.9313725490196079

- Compare the results of your train_test_SVM function between linear, radial basis and polynomial kernel functions.
  - The linear kernel seems to be the best one suited for the task at hand. Since its output values for accuracy, precision and recall are all higher than the output values for the radial basis and polynomial kernels.