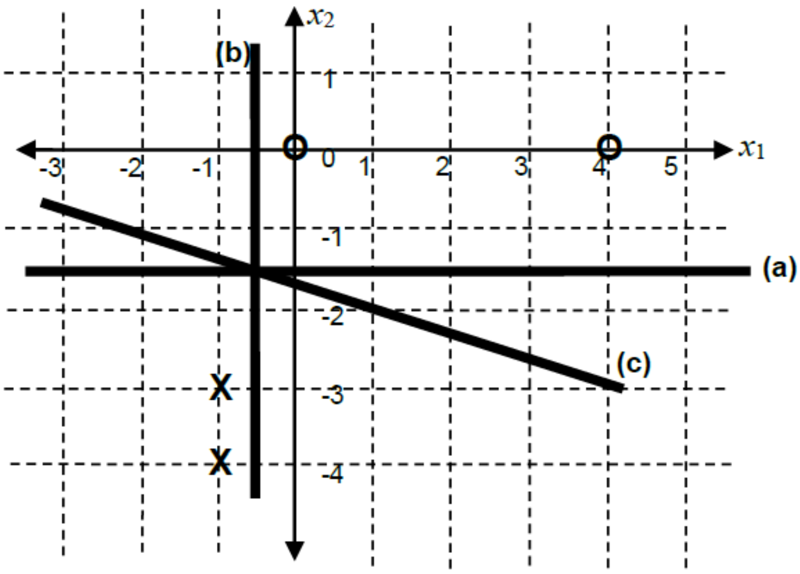
German International University of Applied Sciences
Informatics and Computer Science
Dr. Caroline Sabty
Eng. Nouran Khaled
Eng. Sandra Samuel

Machine Learning, Winter 2024
Solutions Practice Assignment 8

Exercise 8-1

Consider the dataset shown in the figure below where a linear Support Vector Machine (SVM) without slack variables is supposed to be used:



a) **Which** of the decision boundaries (a), (b) or (c) shown on the figure would be the resulting decision boundary of linear SVM? **Show** your calculations. When answering this question, no need to solve by optimizing the SVM objective function.

b) **What** are the support vectors based on your answer in (a)?

c) **How** would adding a training point in location (1, 1) to the dataset that belongs to the (**O**) class change the decision boundary?

Solution:

a) For (a), the margin is 1.5 (half the distance between the lines passing through the nearest points from both classes). For (b), the margin is 0.5. For (c), the margin is 1.5811. Since the SVM is a maximum margin classifier, then the decision boundary is (c).

b) (0, 0) from Class O and (-1, -3) from Class X.

c) It will have no impact since this point is outside the margin beyond the support vectors and the decision boundary of SVM is solely based on the location of the support vectors.

Exercise 8-2

Consider a Support Vector Machine (SVM) without slack variables that is supposed to be used to classify the 1-dimension data given below:

x	t
-2	-1
1	1
3	1

a) Given that the objective function of SVM takes the form
$J = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j t_i t_j \phi(x_i)^T \phi(x_j)$
Find J for the dataset given above in terms of α_i if the kernel used is $K(x_i, x_j) = |x_i||x_j|$. (Don't solve the optimization problem. Just find the expression of J).

b) **Would** solving this optimization problem for this data using the given kernel lead to a decision boundary that classifies the given data correctly? **Explain** your answer.

Solution:

a) In order to get the objective function, we need to find what is $\phi(x)$. Using the kernel function given,
$\phi(x_i)^T \phi(x_j) = K(x_i, x_j) = |x_i||x_j|$

$$K(x_i, x_j) = \begin{bmatrix} |x_1||x_1| & |x_1||x_2| & |x_1||x_3| \\ |x_2||x_1| & |x_2||x_2| & |x_2||x_3| \\ |x_3||x_1| & |x_3||x_2| & |x_3||x_3| \end{bmatrix} = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{bmatrix}$$

$J = \alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2}[\alpha_1\alpha_1(-1)(-1)(4) + \alpha_1\alpha_2(-1)(1)(2) + \alpha_1\alpha_3(-1)(1)(6) + \alpha_2\alpha_1(-1)(1)(2) + \alpha_2\alpha_2(1)(1)(1) + \alpha_2\alpha_3(1)(1)(3) + \alpha_3\alpha_1(-1)(1)(6) + \alpha_3\alpha_2(1)(1)(3) + \alpha_3\alpha_3(1)(1)(9)]$
$= \alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2}[4\alpha_1\alpha_1 - 2\alpha_1\alpha_2 - 6\alpha_1\alpha_3 - 2\alpha_1\alpha_2 + \alpha_2\alpha_2 + 3\alpha_2\alpha_3 - 6\alpha_1\alpha_3 + 3\alpha_2\alpha_3 + 9\alpha_3\alpha_3]$
$\therefore J = \alpha_1 + \alpha_2 + \alpha_3 - \frac{1}{2}[4\alpha_1^2 + \alpha_2^2 + 9\alpha_3^2 - 4\alpha_1\alpha_2 - 12\alpha_1\alpha_3 + 6\alpha_2\alpha_3]$

b) No because using this kernel will change the point (-2) to become +2 and so the data will become non-linearly separable and so it can't be solved by this SVM since it doesn't have slack variables.

Exercise 8-3

Lab Question

You will explore the application of Support Vector Machines (SVM) using the scikit-learn library in Python. Follow the steps outlined below:

- Dataset:
  - Load the Iris dataset from scikit-learn.
  - Split the dataset into training and testing sets.
- Data Preprocessing:
  - Standardize the features.
- SVM Classifier:
  - Create an SVM classifier using a linear kernal with default parameters.
  - Train the classifier on the training set.
- Predicitions and Evaluation.
  - Make predictions on the test set.
  - Evaluate the performance of the classifier using accuracy