Homework 3

Submission Instructions

- Homework is due on: **Sunday 18/08/19 23:55**.
- Homework should be done **only in pairs**. Each pair is to do their own work, separate from the other pairs.
- We prefer you type your submission, however, you may submit scanned handwritten material as long as it is **clear and readable**.
- Submit **only one** PDF file. Please **write your ID** on the top of the file.
- Submission is done via **Moodle** website.
- Homework can be done using either MATLAB or Python.

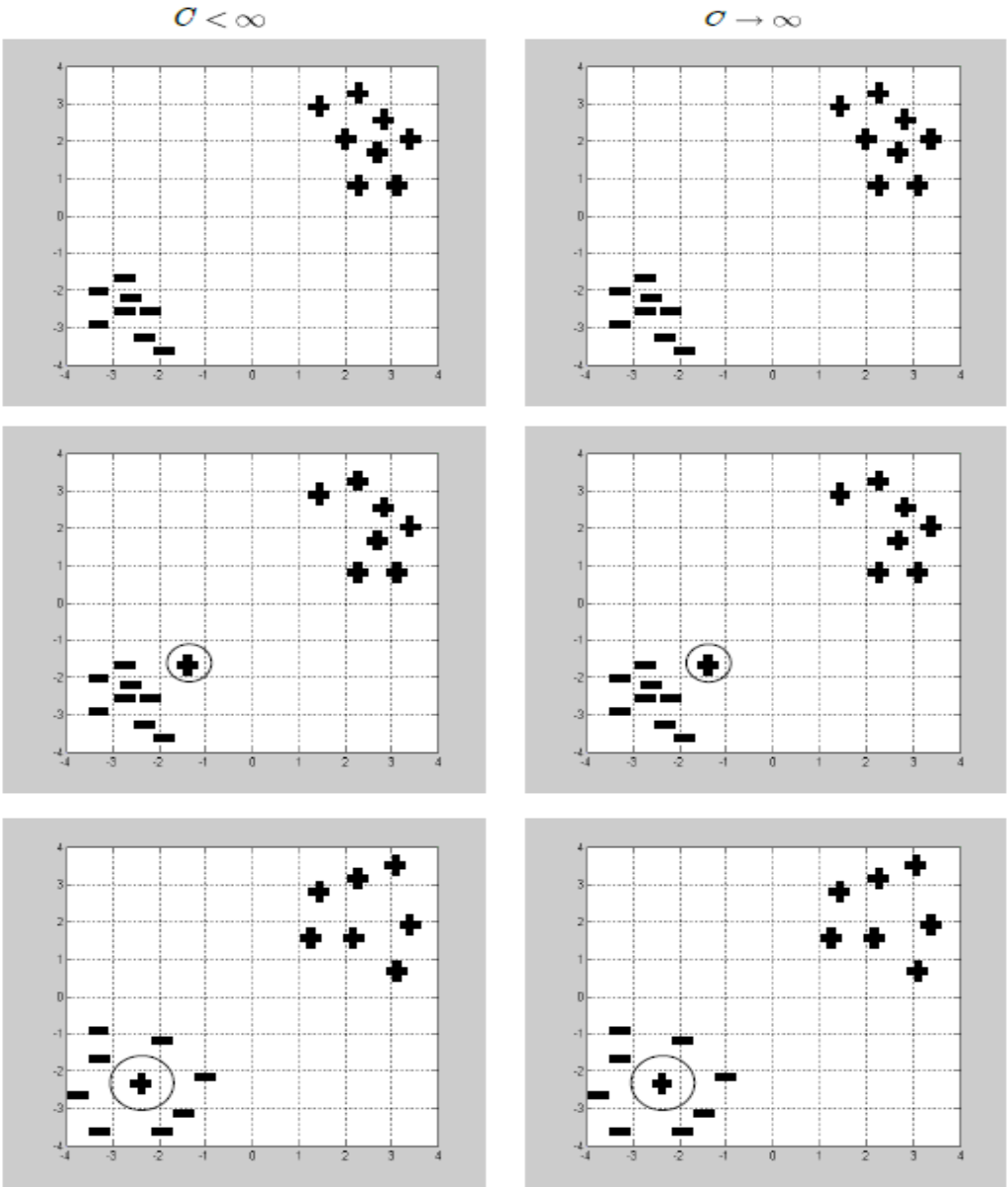
Question 1 - Soft SVM

(a) Sketch the separating hyperplane for the three datasets below and for two values of C:

- In the left column sketch the hyperplane for the case $C \to \infty$.
- In the right column sketch the hyperplane for the case $C < \infty$.
  If the separation hyperplane does not exist, explain why.

(b) In the last two problems (4 last figures) there is a circled data point, what is the suitable value of ξ (Equal to 0, between 0 to 1, greater than 1) for that point? Explain.

You should attach this page to your homework.



Question 2

Consider a training set $\{x_i\}_{i=1}^n$, $(x_i) \in \mathbb{R}^n$ with labels $y_i \in \{0, 1\}$ (binary problem). After training a SVM classifier with $C \to \infty$, the number of support vector received was $k = 2$, $(k < n)$. Later, a new example x_{n+1} was added to the training set and a new classifier was learned. Determine which of the following options are possible, there could be more than one possible option (It is recommended to explain with a sketch):

(a) The number of support vector remained $k = 2$.

(b) The number of support vector grew to $k + 1$

(c) The number of support vector grew to $n + 1$.

Question 3

Consider two kernel functions $k_1, k_2 : X \times X \to \mathbb{R}$. t is known that the classification problem is linearly separable for k_1 but not for k_2. We define a new kernel function as

$$k_3(x, x') = k_1(x, x') + k_2(x, x').$$

(a) Is $k_3(x, x')$ is a valid kernel function? If yes, then explicitly show that it satisfies the conditions required from a kernel function.

(b) Is the classification problem is linearly separable for $k_3(x, x')$?

Question 4

Which of the classifiers below have a zero training error on the following dataset:

X	Y
(-1,-1)	-1
(-1,+1)	+1
(+1,-1)	+1
(+1,+1)	-1

1. Linear SVM.
2. SVM with a polynomial kernel function of degree 2.
3. SVM with a Gaussian kernel function $K_\lambda(x, z) = e^{-\frac{||x-z||^2}{\lambda}}$.

Question 5

Consider the following function:

$$f(x,y) = -20\left(\frac{x}{2} - x^2 - y^2\right)\exp\left(-x^2 - y^2\right).$$

(a) Plot this function in the range of $-3 \leq x, y \leq 3$ (You may use MATLAB functions *mesh* and *meshgrid*).

(b) Implement the gradient descent method for finding the minimum point. Attach your code to your submitted pdf file.

(c) Initialize your algorithm with the following values:

- $[x_0, y_0] = [0.1, 1]$.
- $\eta = 0.01$.

Plot the convergence graph of the algorithm (i.e. the value of the function at each step). To what point if any the algorithm converges?

(d) Initialize your algorithm with the following values:

- $[x_0, y_0] = [1.5, -1]$.
- $\eta = 0.05$.

Plot the convergence graph of the algorithm. To what point if any the algorithm converges? Which phenomenon can be observed?

(e) Initialize your algorithm with the following values:

- $[x_0, y_0] = [1.5, -1]$.
- $\eta = 0.01$.

Plot the convergence graph of the algorithm. To what point if any the algorithm converges? Compare your results with the results of part (d).