Problem 1

Suppose we use the soft-margin linear SVM on a data set $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^m$:

$$\begin{aligned}\min_{\boldsymbol{w}, b, \boldsymbol{\xi}} \ & \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} + C\sum_{i=1}^m \xi_i \\ \text{s.t.} \ & y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1 - \xi_i, \quad i = 1, 2, \ldots, m, \\ & \xi_i \geq 0, \quad i = 1, 2, \ldots, m.\end{aligned} \tag{1-1}$$

Please answer the following questions:

(1) Where does a data point lie relative to where the margin and decision boundary are when $\xi_i = 0$? Is this data point classified correctly?

(2) Where does a data point lie relative to where the margin and decision boundary are when $0 < \xi_i \leq 1$? Is this data point classified correctly?

(3) Where does a data point lie relative to where the margin and decision boundary are when $\xi_i > 1$? Is this data point classified correctly?

Problem 2

Recall the primal problem of hard-margin linear SVM is

$$\begin{aligned}\min_{\boldsymbol{w}, b} \ & \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} \\ \text{s.t.} \ & y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \geq 1_i, \quad i = 1, 2, \ldots, m.\end{aligned} \tag{2-1}$$

The linear classifier (2-2) can be induced by solving the dual of this primal problem to perform linear classification. Moreover, we can replace $\boldsymbol{x}_i^\top \boldsymbol{x}$ with a kernel $K(\boldsymbol{x}_i, \boldsymbol{x})$ to achieve a non-linear classifier.

$$f(\boldsymbol{x}) = \sum_{i=1}^m \alpha_i y_i (\boldsymbol{x}_i^\top \boldsymbol{x}) + b \tag{2-2}$$

In Figure 1, there are 6 different SVMs with different patterns of decision boundaries. The training data is labeled as $y \in \{-1, 1\}$, represented as the shape of circles and squares respectively. Support vectors are drawn in SOLID circles. Label each plot in Figure 1 with the letter of the optimization problem below and explain WHY you pick the figure for a given kernel. (Note that one of the plots does not match to anything.)

1. A soft-margin linear SVM with $C = 0.02$.
2. A soft-margin linear SVM with $C = 20$.
3. A hard-margin kernel SVM with $K(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^\top \boldsymbol{v} + (\boldsymbol{u}^\top \boldsymbol{v})^2$.
4. A hard-margin kernel SVM with $K(\boldsymbol{u}, \boldsymbol{v}) = \exp(-5\|\boldsymbol{u} - \boldsymbol{v}\|^2)$.
5. A hard-margin kernel SVM with $K(\boldsymbol{u}, \boldsymbol{v}) = \exp(-\frac{1}{5}\|\boldsymbol{u} - \boldsymbol{v}\|^2)$.

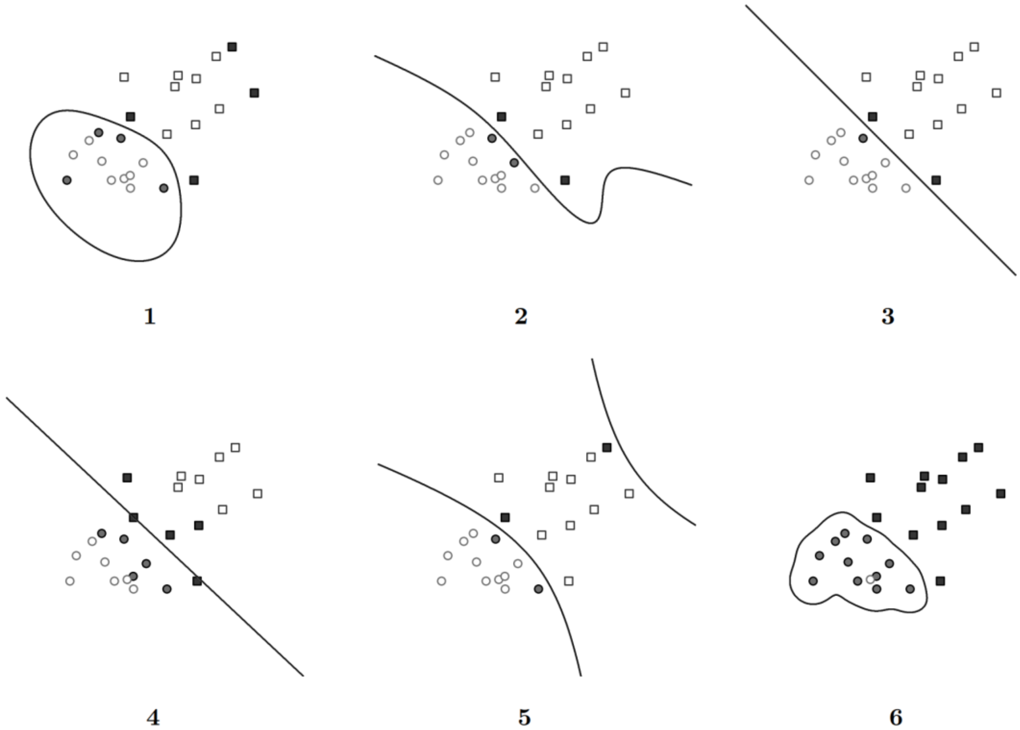

Figure 1: SVM Decision Boundaries

Problem 3

Consider the following non-linear optimization problem:

$$
\begin{aligned}
\min \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & x_1 + x_2 = 1 \\
& x_2 \le \alpha.
\end{aligned}
\tag{3-1}
$$

Here, $x_1,\ x_2,\ \alpha \in \mathbb{R}$.

(1) Write the KKT conditions for the problem.

(2) Find the optimal solution of the problem.