CS5691 - Pattern Recognition and Machine Learning
Jul – Nov, 2025

"Practice Worksheet: SVM"

November 7, 2025

1. Consider a SVM Hard Margin problem where the decision boundary is defined by $z(x) = 0$ where $z(x) := w^T x + b$.
   (i) Derive an expression for the (Euclidean) distance between the margins (margin boundaries).
   (ii) What is the distance of the origin $(0, 0)$ to the decision boundary $z(x) = 0$?

2. Consider a soft margin SVM where C is the penalty parameter. Explain how the behavior of SVM as a classifier will change as C is increased from a very small value to a very high value.

3. Let $u \in \mathbf{R}^d$ be a point. Let $w \in \mathbf{R}^d, b \in \mathbf{R}$ and the hyperplane given by w, b is $\{x \in \mathbf{R}^d : w^T x + b = 0\}$. Consider the following problem of projection of the point u on to a (hyper)plane given by w, b.

$$\min_{v \in \mathbf{R}^d} \frac{1}{2}\|v - u\|^2$$

$$\text{s.t. } w^T v + b = 0$$

   Derive the solution to the above problem via solving the Lagrangian dual (which is an unconstrained quadratic problem, and hence can be easily solved; reviewing the minimax theorem and KKT conditions seen in class may help). Then show that the distance of the point u to the hyperplane given by w, b is $\dfrac{|w^T u + b|}{\|w\|}$.

4. Let $\{(x_1, y_1), ..., (x_n, y_n)\}$ be a linearly separable binary classification dataset. Let w^*, b^* be any solution to the problem below:

$$\max_{w \in \mathbb{R}^d, b \in \mathbb{R}} \frac{1}{\|w\|}$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1$$

   Show that $\min_{i \in [n]} y_i(w^T x_i + b) = 1$.

5. Consider a soft margin SVM problem with C set to some constant. Let α^* be the dual solution, and let w^*, b^* be the primal solution. Let the dataset be (x_i, y_i) with i ranging from 1 to n.
   (i) If $\alpha^* = 0$, what are the possible range of values of $(w^*)^T x_i + b^*$?
   (ii) If $0 < \alpha^* < C$, what are the possible range of values of $(w^*)^T x_i + b^*$?
   (iii) If $\alpha^* = C$, what are the possible range of values of $(w^*)^T x_i + b^*$?

   (Hint: Use KKT complementary slack conditions and $\beta_i^* = C - \alpha_i^*$)

6. Consider the following 1-dimensional classification dataset with 8 points given by:

$$X^T = \begin{bmatrix} 1 & 2 & 4 & 5 & 6 & 7 & 9 & 10 \end{bmatrix}$$

$$y^T = \begin{bmatrix} +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \end{bmatrix}$$

(i) Solve Kernel SVM (Hard Margin) (i.e. give the optimal α^*) with $K(u, v) = exp(-\gamma(u - v)^2)$ where $\gamma = 0.1$.
Hint: $\alpha_2^* = \alpha_7^* = \alpha_3^* = \alpha_6^* > 0$ while all other $\alpha_i^* = 0$.
Find such an α^* and then use KKT conditions to show that it is optimal.

(ii) Give b^* for all α^* above.

(iii) Give the decision function $(w^*)^T\phi(x) + b^*$ for $x \in \mathbf{R}$.

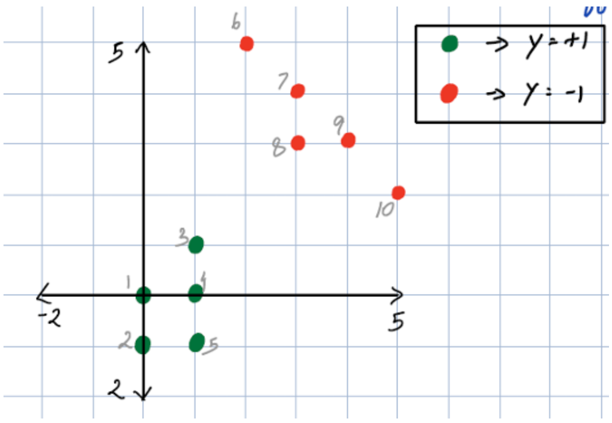
7. Consider the following hard margin SVM problem with both w and b. Assume the kernel to be the linear kernel.

(i) Argue what points are support vectors.

(ii) Argue what would be the optimal hyperplane and give w^*, b^*.

(iii) Also argue what α^* should be.

(You can use software to check intuition, and use KKT conditions to verify if a proposed α^* is actually an optimal solution.)



(iv) Repeat parts (i), (ii), (iii) if point (x_8, y_8) is removed.

(v) Repeat parts (i), (ii), (iii) with point $(x_8, y_8), (x_9, y_9), (x_7, x_7)$ removed. (Hint: Optimal α^* need not be unique even if w^* and b^* are.)

8. Consider the following 2-dimensional classification dataset with 5 points given by:

$$X^T = \begin{bmatrix} 1 & 1 & 2 & 4 & 5 \\ 1 & 0 & 5 & 4 & 2 \end{bmatrix}$$

$$y^T = \begin{bmatrix} +1 & +1 & -1 & -1 & -1 \end{bmatrix}$$

Consider the hard margin SVM problem with linear kernel $k(u, v) = u^T v$.

(i) Give the support vectors just by looking at the data. Give reasons.

(ii) Give the dual solution α^* using the answer to the above part.

(iii) Check if the entire solution got above is the right answer using KKT conditions. (Thus also checking the first part guessed by "eyeballing".)

(iv) Derive the primal solution w^*, b^* from the dual solution α^* and draw a figure illustrating the final solution.

9. Consider the following 2-dimensional binary classification dataset with 10 points given by

$$X^T = \begin{bmatrix} 1 & 1 & 2 & 2 & 4 & 4 & 5 & 5 & 2.9 & 3.1 \\ 0 & 1 & 0 & 1 & 3 & 4 & 3 & 4 & 6 & 6 \end{bmatrix}$$

$$y^T = \begin{bmatrix} -1 & -1 & -1 & -1 & +1 & +1 & +1 & +1 & -1 & +1 \end{bmatrix}$$

Consider the soft-margin linear SVM problem with $C = 0.1, 1, 10, 100$. For each C evaluate the following w, b. By evaluate, we mean you should give the slack variable ξ that make the w, b, ξ feasible, and also give the value of the objective.
(i) $w = (\frac{1}{2}, 0), b = \frac{-3}{2}$ (ii) $w = (1, 0), b = -3$ (iii) $w = (4, 0), b = -12$ (iv) $w = (16, 0), b = -48$
(v) $w = (64, 0), b = -192$ (vi) $w = (\frac{1}{4}, \frac{1}{4}), b = \frac{-5}{4}$ (vii) $w = (\frac{1}{2}, \frac{1}{2}), b = \frac{-5}{2}$ (viii) $w = (1, 1), b = -5$
(ix) $w = (2, 2), b = -10$ (x) $w = (4, 4), b = -20$