Homework 4 Support Vector Machines

CS461 2025-11-10

Name:____

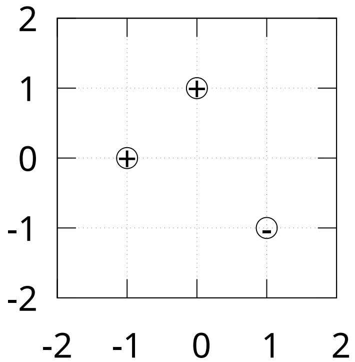
The complexity of support vector machines lies in locating the support vectors. The remainder of the calculations are straightforward (in most cases), as you will see in this homework.

First, remember that we choose alpha to maximize an equation:

$$\underset{\alpha}{\text{argmax}} \quad -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j + \sum_{i=1}^{n}\alpha_i$$

In addition, we have the following two constraints:
$\sum_{i=1}^{N}\alpha_i y_i = 0$
$0 \leq \alpha_i \forall i \in N$

Part 1: Given $X = [(0,1), (-1,0), (1,-1)]$ and $Y = (1,1,-1)$, solve for $\alpha_1, \alpha_2, \alpha_3$. Note that the two points from class 1 are symmetrical relative to the third point, so $\alpha_1 = \alpha_2$.



Part 2: With the values from part 1, solve for w.

Part 3: Solve for w_0. It can be derived from any support vector with the equation $w_0 = y_i - w^T x_i$. In practice, we average over all support vectors because of numerical imprecision, but in this case you can solve using multiple support vectors to verify that your previous solutions were correct.

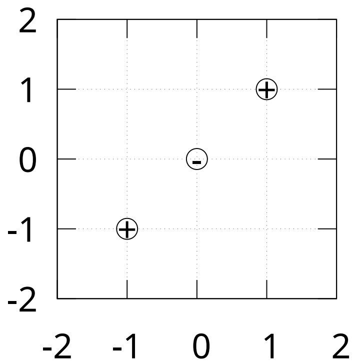
Part 4: $w^T(x, y) + w_0 = 0$ describes the decision boundary, with the two components of w projecting into the x and y axes, respectively. Convert that into the familiar line format, $y = mx + b$. Think of this as multiplying w by (x, y) to indicate its components and then isolate y.

Part 5: Solve for the size of the margin. This is just $m = \frac{1}{\|w\|}$. Your solution can be verified graphically (see the plot in part 1).

Problem 2: Solving with a kernel function is slightly different, as w remains a function of α and is not solved for directly. Solving for α remains the same though, but now the equation is

$$\underset{\alpha}{\text{argmax}} \quad -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j k(x_i, x_j) + \sum_{i=1}^{n}\alpha_i$$

Part 2.1: Given the polynomial kernel $k(x_1, x_2) = (x_1 \cdot x_2 + 1)^2$, $X = [(-1,-1),(1,1),(0,0)]$ and $Y = (1, 1, -1)$, solve for $\alpha_1, \alpha_2, \alpha_3$. Note that the two points from class 1 are symmetrical relative to the third point, so again $\alpha_1 = \alpha_2$. As always, $\sum_{i=1}^{N}\alpha_i y_i = 0$.



Part 2.2: Using those α values, solve for w_0. The solution is still $w_0 = 1 - k(w, x_i)$, but application of the kernel now requires evaluating the function for each support vector. Don't forget to multiply the result of the kernel by $y_i\alpha_i$ for each support vector.