1 Problem Set for In-class Problem Solving

Answer with sufficient details with insightful discussions (not just factual answers) and submit by 11am. (details discussed in the last two lecture sessions)

1. Consider a linearly separable data

$$(-1, +1), (0, -1), (+1, -1)$$

   (a) Write the primal problem. Geometrically show the feasible region in a w, b 2D plane. Show the optimal w as -2 and b as -1. Geometrically validate the solution.

   (b) Write the dual objective $J(\alpha_1, \alpha_2, \alpha_3)$. And the constraint $\sum_i \alpha_i y_i = 0$. Show that the solution is $\alpha_1 = \alpha_2 = 2$ and $\alpha_3 = 0$. Which are the support vectors then?

   (c) Show that the dual solution also leads to the same w and b.

2. Consider a linearly in-separable 1D data

$$(-1, +1), (0, -1), (+1, +1)$$

   Demonstrate how the problem can be solved with a quadratic kernel $(p^T q + 1)^2$

   (a) Write the dual objective with the constraint $\sum_i \alpha_i y_i = 0$

   (b) Substitute for α_2 from the constraint in the objective. (i.e., eliminate α_2)

   (c) Differentiate wrt α_1 and α_3 and equate to zero. Solve to obtain $\alpha_1 = \alpha_3 = 1$ and $\alpha_2 = 2$

   (d) For a new sample the decision is

$$sign(\sum_i \alpha_i y_i \kappa(x_i, x) + b)$$

   Simplify and get this expression as $2x^2 - 1$. (Assume $b = -1$. Indeed, you can find b yourself.)

3. Consider the ExOR Problem:

x_1	x_2	y
-1	-1	-1
-1	+1	+1
+1	-1	+1
+1	+1	-1

   (a) Consider a kernel $(p^T q + 1)^2$ and construct the Kernel matrix

   (b) Write the dual objective in terms of α_1, α_2, α_3 and α_4.

   (c) Differentiate wrt α_1, α_2, α_3 and α_4 and equate to zero. You get four equations.

   (d) Solve (guess) the values for α_1, α_2, α_3 and α_4. Same?

   (e) We know $\phi() = [1, x_1^2, \sqrt{2}x_1x_2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2]^T$. Find w as $\sum_i \alpha_i y_i \phi(x_i)]$. Does it solve the exor? why?

4. VC dimension of a function class is the largest set that can be shattered by a member of the function class.

   (a) Interval $[\alpha_1, \alpha_2]$ on real line. (i.e., +ve is inside $[\alpha_1, \alpha_2]$ else negative)

   (b) Axis parallel rectangles in 2D. (i.e., positive inside and negative outside)

   (c) Convex Polygons in 2D (inside positive and outside negative)

   (d) Union of intervals in 1D.

5. Starting from the Lagrangian

$$L(\mathbf{w}, b, \alpha, \xi) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + \frac{C}{2}\sum_{i=1}^{N} \xi_i^2 - \sum_{i=1}^{N} \alpha_i [y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 + \xi_i]$$

   Derive the dual function for L2 SVM