Quiz 2

Wednesday, April 6, 2022 1:55 PM

- **This is an open-notes, open-book quiz.**
- **You are not allowed to discuss the questions with any individual at any time.**
- **You have 90 minutes to complete and upload your solution file.**
- **You must upload a single file.**
- **You must name your file Lastname_Firstname_Quiz2.***
- **Failing to follow these instructions will result in a grade of zero.**

1. Consider the following classification problem.

$$\underline{x}^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, y^{(1)} = 1; \quad \underline{x}^{(2)} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, y^{(2)} = 1; \quad \underline{x}^{(3)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, y^{(3)} = -1$$

   a. What is the hard-margin SVM solution (weight vector and bias)? Justify your answer.
   b. Compute the size of the margin for the solution in part (a).
   c. Give the set of equations that must be solved to find the λ's.
   d. Would adding the point $\underline{x}^{(4)} = [-2 \;\; 0]^T$, $y^{(4)} = 1$ change the solution (weight vector and bias) or the λ's? Why?

2. Consider the following constrained optimization problem.

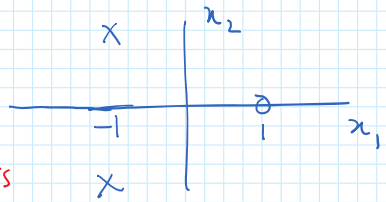
$$\min_x \; x^3 - 6x^2 + 11x - 6$$
$$s.t. \quad 2 \le x \le 3$$

   a. Give the corresponding Lagrangian function.
   b. Give the corresponding KKT conditions.
   c. Are the constraints active? Why?

3. How does the Naïve Bayes classifier address the problem of curse of dimensionality?

(1a) $d(\underline{x}) = -x_1 = 0$

$\Rightarrow \underline{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad w_{d+1} = 0$ 4 pts



(b) $\text{Margin} = \frac{2}{\|\underline{w}\|} = \frac{2}{1} = 2$ 2 pts

(c) $\lambda_1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} - \lambda_3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ } 3 eqn.'s 3 unknowns 6 pts.

$\lambda_1 + \lambda_2 - \lambda_3 = 0$

(d) It will not change $\underline{w}$ or w_{d+1} but λ's may change 2 pts.

2(a) $\mathcal{L}(x, \underline{\lambda}) = x^3 - 6x^2 + 11x - 6 - [\lambda_1 (x-2) + \lambda_2 (-x+3)]$ 2 pts.

(b) $\frac{\partial \mathcal{L}}{\partial x} = 3x^2 - 12x + 11 - \lambda_1 + \lambda_2 = 0$ }

(b) $\frac{\partial L}{\partial x} = 3x^2 - 12x + 11 - \lambda_1 + \lambda_2 = 0$

$\lambda_1 \geq 0 \quad \lambda_2 \geq 0$

$\lambda_1 (x-2) = 0$

$\lambda_2 (-x+3) = 0$

} 4 pts.

(c) no, they are not — soln. of above 4 equations will result in $\lambda_1 = \lambda_2 = 0$ 2 pts.

3) it requires parameter estimation for ℓ, 1-D pdf's (for each class), so ℓN data points would be needed instead of N^ℓ. 6 pts.

30 pts. Max
start with 2 pts.