CSE 326/426 (Fall 2022) Homework 4

Due on 11:55pm, Fri, Sep 23, 2022

Grading: All questions have the same points (10 each). Two questions from the first four will be randomly selected for grading for all students. The last question is required for graduates.

Submitting: Only electronic submissions in PDF format on Coursesite are accepted. Name your file as

<Your LIN>_HW4.pdf

A few ways to create a single PDF file. i) Use Microsoft Word and insert your writing or any figure taken, scanned, or plotted. Then choose "File→Print" in the main menu and you will find an option for outputting the file to a PDF file. ii) Use Latex to write your solution and include any figures. iii) Use the online Google Doc as an alternative of Word. It has sufficient features to combine multiple images and texts. Exporting to PDF files is similar as with Word.

Please DO NOT compress the PDF file as that will slow down the processing of your submission and the grading.

Questions:

1. Given the two training examples $\mathbf{x}^{(1)} = [2, 1]^\top, y^{(1)} = 1$ and $\mathbf{x}^{(2)} = [1, -1]^\top, y^{(2)} = -1$, first write down the SVM primal problem and plug in the data (so that $\mathbf{w}$ and b are variables). Then construct the Lagrangian function $L(\mathbf{w}, b, \boldsymbol{\lambda})$.

2. Write down the KKT conditions of the above problem. Simplify the partial derivatives and write down the elements of the vectors (i.e., don't vectorize the conditions).

3. Continuing the above problem. Write down the dual SVM optimization problem with linear kernel. (*Hint: No need to derive the dual problem, but just plug the data in the dual problem in the lecture note*).

4. Continue the above question. Given the dual parameters $\boldsymbol{\alpha} = [0.5, 1]$, which may not be optimal, recover the primal parameter $\mathbf{w}$ (using the first-order condition of the KKT condition of SVM). Use linear kernel and solve for b by noting that any support vector on the margin should have $y^{(i)}(\mathbf{w}^\top \mathbf{x}^{(i)} + b) = 1$. Then use the dual parameters to classify the test example $\mathbf{x} = [1, 1]^\top$ by computing the classifier output $h(\mathbf{x}; \boldsymbol{\alpha})$.

5. (Graduate only) Prove that there is a function $\psi([x_1, x_2])$, that maps from the input space $\mathbb{R}^2$ to a higher dimensional space $\mathbb{R}^n$, $n > 2$, so that the polynomial kernel $\mathtt{k}(\mathbf{x}, \mathbf{z}) = (\langle \mathbf{x}, \mathbf{z} \rangle + 1)^2$ can be written as $\langle \psi(\mathbf{x}), \psi(\mathbf{z}) \rangle$.

   [*Hints: refer to PRML Eq. (6.12) for help and write down the formula of the function* ψ.]