EE 559 **Homework 3 (Week 5)** Posted: Fri., 2/11/2022
Jenkins Due: Fri., 2/18/2022, 11:59 PM PST

1. You are given the following training data in a 1D (1 feature), 2-class problem:

$$S_1\text{: } x_1 = 0,\ x_2 = 2,\ x_3 = 3$$

$$S_2\text{: } x_4 = -2,\ x_5 = -3,\ x_6 = -4$$

(a) (i) Plot the data in non-augmented feature space. Draw a linear decision boundary H (a point) that correctly classifies them, showing which side is positive.

(ii) Give the weight vector $\underline{w}$ that corresponds to your H. Because the length of $\underline{w}$ is not determined by the given quantities, choose $\underline{w}$ so that $||\underline{w}|| = 1$.

Hint: the boundary is given by $g(x) = 0$, and the decision region for Γ_1 is $g(x) > 0$.

(b) Plot the data points (in augmented form) in augmented feature space. Also show the decision boundary and which side is positive.

(c) Plot the augmented data points as lines $g_{\underline{x}_n}(\underline{w})$ in weight space. Show the positive side of each such line with a small arrow. Show the solution region.

(d) Plot the weight vector $\underline{w}$ of H from part (a) as a point in weight space. Is $\underline{w}$ in the solution region?

For parts (e)-(f) below, you will use the same dataset of points $\underline{x}_n$ given above, except in augmented form as in (b), but you will reflect them and use the reflected data points $z_n\underline{x}_n$ for your plots.

(e) Plot the reflected data points $z_n\underline{x}_n$ in augmented feature space. Draw the linear decision boundary H from your part (b) and its weight vector. Does H correctly classify all the data points?

Hint: write down the general condition for correct classification of reflected data points.

(f) Plot the data points $z_n\underline{x}_n$ as lines in 2D weight space, showing the positive side of each with a small arrow. Show the final solution region.

2. This problem uses the notation we used in Lecture 6, and m is a positive integer. For the following computational complexity:

$$p(m) = 2m + 1000$$

(a) Is $p(m) = O(m)$?

If yes, prove your answer by letting $a = 1$, and solve for what m_0 we have $m \geq p(m)\ \ \forall m \geq m_0$. If you need a larger a, then state what value of a will work. Find the smallest such integer m_0 for the value of a you used.

If no, justify why not.

(b) Is $p(m) = \Omega(m)$?

If yes, prove your answer by letting $b = 1$, and solve for what m_1 we have $m \le p(m) \quad \forall m \ge m_1$. If you need a smaller *b*, then state what value of *b* will work. Find the smallest such integer m_1 for the value of *b* you used.

If no, justify why not.

(c) Is $p(m) = \Theta(m)$?

Justify your answer.

3. This problem also uses the notation of Lecture 6.

(a) Suppose we have a function $p(m)$ that can be expressed as:

$$p(m) = p_1(m) + p_2(m) + p_3(m)$$

and we have:

$$p_k(m) = O(q_k(m)), \quad k = 1,2,3 \qquad \text{(i)}$$

Prove that:

$$p(m) = O(q_1(m) + q_2(m) + q_3(m)) \qquad \text{(ii)}$$

Hints: (i) Use the definition of big-O.

(ii) If you find the problem statement unclear or confusing, try looking at the example in the appendix below.

(b) Is a similar statement to (a) true for $\Omega(..)$? (That is, if you replace each $O(..)$ in part (a) with $\Omega(..)$, would the last equation be true?) Justify your answer.

4. Consider the following computational complexity:

$$p(m) = m^2 \log_2 m + \frac{2m^3}{\log_2 m} + (\log_2 m)^3$$

(a) Is $p(m) = O(m^3)$? If yes, justify it by showing it satisfies the definition. If no, justify why not.

Hint: $p(m)$ is the sum of 3 terms. Try setting $a = 1$, and check each term individually to see if it is $O(m^3)$ (if it is, then give the value of each m_0). Then use the result of Problem 3(a).

(b) Is $p(m) = \Omega(m^3)$? Justify your answer.

5. For a nearest-means classifier, with $C=2$ classes (held constant), $\frac{1}{2}N$ data points in each class (assume N is even), and D features, find the computational time complexity and space complexity, for the operations given in (a) and (b) below.

For all parts of this problem, assume a completely serial computer. Your answer may be in the form of a big-O upper bound; express the big-O bound in simplest form, without making it looser (higher) than necessary. As part of your answer to each part, please:

(i) Give a brief statement of the algorithm you are analyzing

(ii) Show your reasoning in calculating the complexity

(a) Computing each mean vector from the training data (the training phase)

(b) Classifying M data points, given the mean vectors (the classification phase)

(c), (d) Repeat parts (a) and (b), except for a C-class problem, with $\frac{N}{C}$ data points per class, in which C is a variable.

Hint for (d): you may use without proof that finding the minimum of C values that are unsorted can be done in $O(C)$ time and $O(1)$ space.

Appendix - Example (relates to Problem 3)

Suppose we want to find and prove the (tightest) asymptotic upper bound of $p(m)$, with:

$$p(m) = 3m^3 + 100m^2 \log_2 m + 0.1(2^m)$$

Applying the definition directly to $p(m)$ (especially to prove your bound, including finding m_0) might be difficult. Instead, you could use the result of Problem 3a, to apply the big-O bound to each term independently:

$$3m^3 = O(m^3)$$
$$100m^2 \log_2 m = O(m^2 \log m)$$
$$0.1(2^m) = O(2^m)$$

Then using Problem 3a equation (ii), we can conclude:

$$\begin{aligned} 3m^3 + 100m^2 \log_2 m + 0.1(2^m) &= O(m^3 + m^2 \log m + 2^m) \\ &= O(2^m) \end{aligned}$$