1 EM Part 1

A professor proposes the following scheme for assigning grades: He will use the EM algorithm to identify 5 normally distributed clusters of students based upon their numerical grades. He will then sort the clusters by their means and assign grades: A, B, C, D and F from highest to lowest. Assume that the professor is making a good faith effort to assign fair grades, then carefully state all of the assumptions that the professor’s proposal is making about the distribution of grades.

2 EM Part 2

Implement code to generate samples according to the professor’s assumptions and code to determine the clusters and assign grades. Show some examples of how your code works by providing some sorted list of sample numerical scores and letter grade assignments. If the data obey the professor’s assumptions, do you think his method is fair? Justify your answer

3 EM Part 3

Generate some data that violate the professor’s assumptions and show (by running your code) how this can lead to an extremely unfair outcome.

4 Perceptrons

Provide an algorithm that takes as input the dimensions of a hyperrectangle and produces the weights for a multilayer perceptron that checks if its inputs lie within this hyperrectangle.

5 Understanding Multilayer Networks

Read the excerpt from Churchland and describe two distinct areas where his presentation is misleading our plain wrong.

6 Local optima

Provide a rigorous argument that the expressive power of multilayer neural networks (with differentiable activation functions) implies the possibility of the existence of local optima in the error function. (Your argument does not need to apply to all target functions.)