Expectation Maximization

Week 4 - Session 1

Problem

Imagine a machine learning class where the probability that a student gets an "A" grade is P (A) = 1/2, a "B" grade P (B) = μ , a "C" grade P (C) = 2μ, and a "D" grade P (D) = 1/2-3μ. We are told that c students get a "C" and d students get a "D". We don't know how many students got exactly an "A" or exactly a "B". But we do know that h students got either an A or B. Therefore, a and b are unknown values where a + b = h. Our goal is to use expectation maximization to obtain a maximum likelihood estimate of μ.

Solution:

1. Expectation step: Which formulas compute the expected values of a and b given μ?

$$\frac{a}{h} = \frac{1/2}{1/2+\mu}, \frac{b}{h} = \frac{\mu}{1/2+\mu}$$

$$\therefore \hat{a} = \frac{h}{1+2\mu}, \hat{b} = \frac{h\mu}{1+2\mu}$$

2. Maximization step: Given the expected values of a and b which formula computes the maximum likelihood estimate of μ? Hint: Compute the MLE of μ assuming unobserved variables are replaced by their expectation.

$$Log(P) = a * log(1/2) + b * log(\mu) + c * log(2\mu) + d * log(1/2 - 3\mu)$$

To maximize Log(P), we should make its derivation 0.

$$\frac{b}{\mu} + \frac{2c}{2\mu} + \frac{-3d}{1/2 - 3\mu} = 0$$

$$\frac{b+c}{\mu} + \frac{-6d}{1-6\mu} = 0$$

$$b + c - 6\mu(b+c+d) = 0$$

$$\therefore \mu = \frac{b+c}{6(b+c+d)} = \frac{h}{6(h+d)}$$