Vishwanath Dattatreya Doddamani A0286188L e1237250@u.nus.edu

CS5340 ASSIGNMENT 2 – Part 2 REPORT

Parameter Learning

Details of function implementation:

➔ **Derivations:**

$$L = argmax_{\theta_u} \sum_{n=1}^{N} -\{\frac{1}{2} log(2\pi\sigma_u^2) - \frac{1}{2\sigma_u^2}(x_{u,n} - (\sum_{c \in x_{\pi_u}} w_{uc} x_{uc,n} + w_{u0}))^2\}$$

Differentiating L, w.r.t $w_{u0}, \dots, w_{uC}$, we get the following C+1 equations, where C is the number of parents of a particular node.

$$\frac{\partial L}{\partial w_{u0}} = \sum_{n=1}^{N} (x_{u,n} - (w_{u1}x_{u1,n} + \dots + w_{uC}x_{uC,n} + w_{u0})) = 0$$

$$\frac{\partial L}{\partial w_{u1}} = \sum_{n=1}^{N} (x_{u,n} - (w_{u1}x_{u1,n} + \dots + w_{uC}x_{uC,n} + w_{u0}))x_{u1,n} = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial w_{uC}} = \sum_{n=1}^{N} (x_{u,n} - (w_{u1}x_{u1,n} + \dots + w_{uC}x_{uC,n} + w_{u0}))x_{uC,n} = 0 - (1)$$

Differentiating L w.r.t σ^2, we get the below equation which is used to calculate variance.

$$\frac{\partial L}{\partial \sigma^2} = \frac{\sum_{n=1}^{N} (x_{u,n} - (\sum_{c \in x_{\pi_u}} w_{uc} x_{uc,n} + w_{u0}))^2}{N} - (2)$$

➔ **Functions:**

a. **`_learn_node_parameter_w()`** Using equation (1), the parameters $w_{u0}, \dots, w_{uC}$ are solved. The input array is modified to consist the observations of parents of the node which is passed to np.linalg.solve() along with observations in Ax = b format. The values computed by np.linalg.solve() is then returned.

b. **`_learn_node_parameter_var()`** Using equation (2), the parameter σ^2 is computed. The function implementation for this code is very direct and do not need much modification once the weights are computed. The variance is computed and then returned.

Challenges Faced:

1. Understanding equation (1) is a lot hard than it might seem. However, once it is understood and figured that the equation needs to be modified in accordance with np.linalg.solve(), it is implementable within minutes.

Conclusion:

Part 2 of Assignment 2 makes one realize the previous assumptions (such as parameters which are given by default). This makes one completely understand the problem from framing to end result.