

# CMSE890 Project Proposal

Haiyang Yu

October 8, 2019

## 1. Group Members

Haiyang Yu

## 2. The Program

In statistics, parameters estimation is an important component. Maximum Likelihood Estimation (MLE) is the method to estimate parameters by maximizing the likelihood function. Suppose the parameter is  $\theta$  in parameter space  $\Theta$ . The observed sample  $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$  has the distribution  $f(\mathbf{y}, \theta)$  where parameters  $\theta$  is unknown. To get the best estimation  $\hat{\theta}$ , MLE maximize the likelihood function

$$\hat{\theta} = \arg \max_{\theta \in \Theta} f(\mathbf{y}, \theta)$$

In general, we take the derivatives for the log-likelihood function and let it be 0 to get the analytical solution of  $\hat{\theta}$ .

$$\frac{\partial \log f(\mathbf{y}, \theta)}{\partial \theta} = 0$$

However, some likelihood functions are complicated that we cannot get the analytical solution. In this case, many people use EM algorithm to get the numerical solution of  $\hat{\theta}$ . EM algorithm has 2 steps to get the maximum. The first step is using the initial parameter  $\theta_0$  to calculate the  $Q$  function. The second step is to maximize the  $Q$  function. And enter the next iteration. It has been proved that EM algorithm converges. But if we regard it as an optimization problem, can we apply the optimization methods on the likelihood function, and is it more efficient than EM?

## 3. The Approaches

(1)

Understanding the EM algorithm and some other optimization algorithms like SGD and Adam.

(2)

Understanding the typical model:

I have 3 coins  $A, B$  and  $C$ . First, I toss coin  $A$ . If it is positive side, I toss coin  $B$ , else I toss coin  $C$ , and write down the side types. What has been observed is the side of coin  $B$  and  $C$  like "0, 1, 1, 0, 1, ...". We don't know whether the result is from coin  $B$  or coin  $C$ . Suppose the positive-side probability of coins  $A, B, C$  is  $\pi, p, q$ , then use the observed data to estimate  $\pi, p, q$ .

**(3)**

Apply EM algorithm and other optimization methods on the model I mentioned. Compare different algorithms about convergence and rates.

**(4)**

Analyse the result in (3). Complain why it is different.(If I am able to do.)

#### **4. Task Distribution**

I will do the whole project task.

#### **5. Milestone**

Before Oct 31, learn EM algorithms and code it.

Before Nov 20, learn SGD, Adam and other methods and code it.

Before Dec 9, finish the final report.