PHYS-GA2000-PS7

Ahmet Koral Aykın

November 10, 2023

1 Introduction

In this problem set, we are asked to find the minimum of a specified function. In addition, we are interested in logistic regression which includes minimization of negative log-likelihood.

2 Methods

The algorithm used for minimization is called Brent's method, in which the iterations are performed according to golden section method and the minimum is calculated by finding that of a parabola fitted to triplet. Under certain conditions (which can be found in lecture notes), the bracketing interval is updated using the golden ratio. The function that we want to minimize is as follows

$$f(x) = (x - 0.3)^2 exp(x) \tag{1}$$

The plot of the function around the minimum is represented in Figure 1. Based on that, the first triplet is chosen as $(a, b, c) = (-1, 1.1, 4)$.

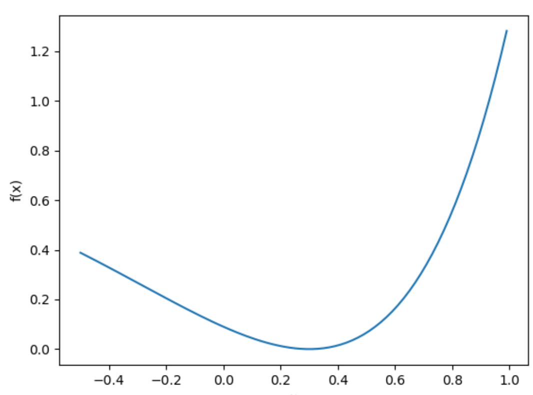

Figure 1: $f(x)$ near the minimum.

For the 2nd problem, we have a logistic function given by

$$p(x) = \frac{1}{1 + exp[-(\beta_0 + \beta_1 x)]} \tag{2}$$

where x is the age of the person under consideration and accordingly $p(x)$ gives the probability of that person knowing the meaning of 'Be kind, rewind'. The answers are encoded as 0s and 1s where 1 corresponds to 'yes' and 0 to 'no'. And our specific goal is to find the optimum parameters for the logistic function given above that describes the data in the best possible way. If we consider a person who is 50 years old, the probability of that person knowing the meaning for different values of logistic function parameters is given in Figure 2.

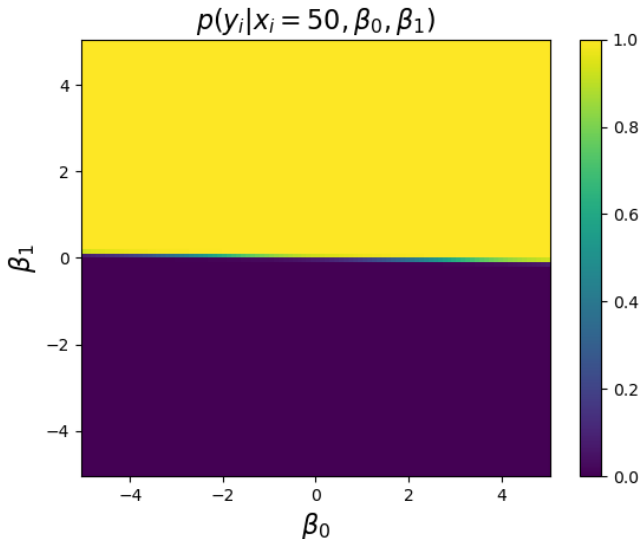

Figure 2: The probability of a person knowing the meaning of the phrase for different values of parameters.

Basically our goal is to find optimum parameters such that the logistic function reflects the actual data. For that purpose, we need first to calculate log-likelihood. The log-likelihood as a function of logistic function parameters is presented in Figure 3. To achieve our goal of finding the optimum parameters, we need to maximize the log-likelihood or minimize the negative log-likelihood. For that purpose, we will use scipy optimization as recommended.

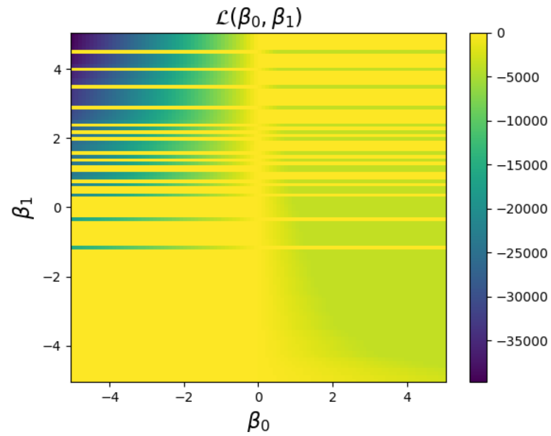

Figure 3: Log-likelihood as a function of the logistic function parameters for the given data set.

3 Results

Figure 4 illustrates how the prediction for the minimum of the function evolves in each parabolic step.

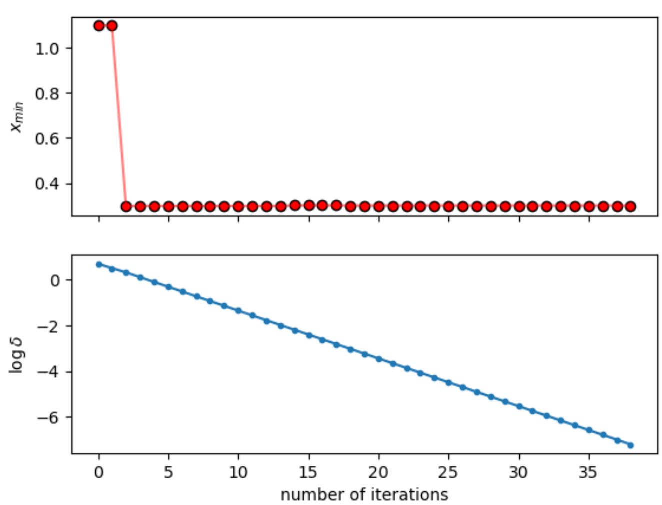

Figure 4: Convergence of the minimum value found from parabolic interpolation as a function of the number of iterations.

The minimum value obtained when bracketing length is lower than the tolerance value is found as 0.2999999995629871. Using scipy.optimize.brent, the minimum of the function is found as 0.300000000023735 which is slightly off what we found using our own algorithm. The difference is due to the initial guess. For scipy method, initial bracket is not specified.

The logistic function with the optimum parameters that maximizes the log-likelihood and the data is presented in Figure 5.

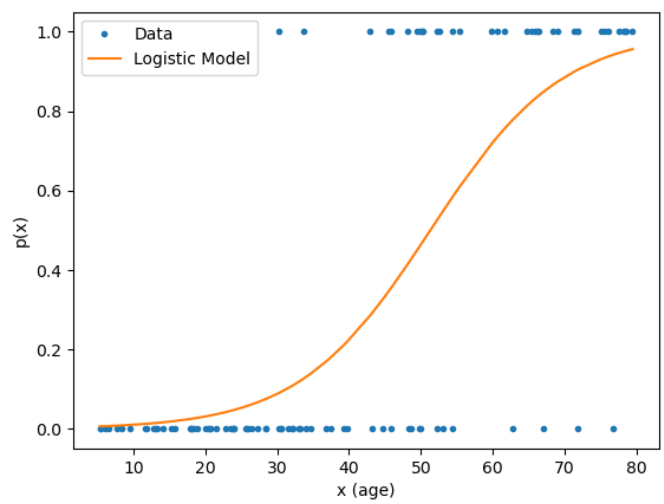

Figure 5: Logistic function with the optimum parameters and the data.

Based on Figure 1, it can be suggested that the logistic model describes the data well. To put it another way, data implies that older people are more likely to know the meaning of the phrase in comparison to young people. So, it is expected that the logistic model converges to 1 (0) as the x (age) increases (decreases). The optimum parameters that maximizes (minimizes) the log-likelihood (negative log-likelihood) are found as following

$$(\beta_0, \beta_1) = (-5.62023229, 0.10956338) \tag{3}$$

Moreover, the associated errors are calculated as follows

$$d(\beta_0, \beta_1) = (0.04895869, 0.00443833) \tag{4}$$

Finally, the covariance matrix of the optimum parameters (β_0, β_1) is calculated as follows

$$C = \begin{pmatrix} 2.39695377 \times 10^{-3} & -7.33380412 \times 10^{-5} \\ -7.33380412 \times 10^{-5} & 1.96988104 \times 10^{-5} \end{pmatrix} \tag{5}$$