THE BULLETIN OF OKAYAMA UNIVERSITY OF SCIENCE No.57 A pp.37-40 (2021)

Improving Fisher's scoring method with two-step iterations for a one-parameter case

Michio Sakakihara and Masahiro Kuroda*
Department of Information Science, Faculty of Informatics,
** Department of Management, Faculty of Management,*
Okayama University of Science,
1-1 Ridai-cho, Kita-ku, Okayama 700-0005, Japan

(Received November 1, 2021; accepted December 9, 2021)

The Fisher's scoring method is one of most popular iterative methods for seeking parameters in density functions from an observed data set. Compared to the Newton method, the computational complexity to evaluate derivative coefficient to proceed the iteration is small, especially when the data size is large. On the other hand, it is well known that the Fisher's scoring method has a slower convergence speed than the Newton method. In this note, we present some improved method via 3rd order iterative methods for univariate non-linear equations. By showing some numerical result for the muon decay problem, it is shown the efficiency of the present modification.

1. Introduction

For the maximal likelihood problem, we employ some iterative method to seek an approximate value of the parameter such as the Newton method, EM algorithm and Fisher's scoring algorithm. The EM algorithm is a robust iterative solution but slow convergence[1). On the other hand, the Newton method is of 2nd order convergence but the convergence characteristics is local. The Fisher's scoring algorithm is widely used a simplified Newton-type iterative method for this problem. The convergence rate of the Fisher's scoring method is less than the Newton method. The Newton method has been modified to improve the convergence rate[2-6)]. It is worth to discuss applying some modifications of the Newton method to the Fisher's scoring iteration. We present three methods in this article. The muon decay, which is a one parameter problem, is utilized to examine the present methods, numerically. We discuss the efficiency of the present methods from numerical results.

2. Fisher's scoring algorithm

Let us explain the Fisher's scoring algorithm for the one parameter case, briefly. The likelihood function is given by

$$L(\alpha) = \prod_{i=1}^{n} f(x_i|\alpha) \tag{2.1}$$

where $f(x|\alpha)$ is a density function. The log-likelihood is obtained such as

$$\ell(\alpha) = \log L(\alpha) \tag{2.2}$$

where $\ell(\alpha)$ is a smooth function. The derivative of the log-likelihood

$$s(\alpha; X) = \frac{d}{d\alpha}\,\ell(\alpha) = \ell'(\alpha) \tag{2.3}$$

is named the score function, in which X is the random observation. The problem is to determine the parameter value to maximize the log-likelihood function. In order to seek the point for the problem, we solve the equation

$$\ell'(\alpha) = 0 \ . \tag{2.4}$$

When we apply the Newton iteration to (2.4), we obtain the iteration

$$\alpha_{k+1} = \alpha_k - \ell'(\alpha_k)/\ell''(\alpha_k) \tag{2.5}$$

where $\ell''(\alpha) = d\,\ell'(\alpha)/d\alpha$.

Taking the expectation of $\ell''(\alpha)$ we have the Fisher information $I(\alpha)$ such as

$$I(\alpha) = -E\big(\ell''(\alpha)\big) = -\int_a^b \ell''(\alpha) f(x|\alpha)dx. \tag{2.6}$$

The Fisher's scoring iteration is expressed as

$$\alpha_{k+1} = \alpha_k + \ell'(\alpha_k)/I(\alpha_k). \tag{2.7}$$

3. Muon decay

In order to illustrate the computational process, we pick up the problem of the muon decay in the text book [1]. The angle θ at in electrons are emitted muon decay has a distribution with the density:

$$f(x|\alpha) = \frac{1+\alpha x}{2}, -1 \le x \le 1 \text{ and } -1 \le \alpha \le 1 \tag{3.1}$$

where $x = \cos\theta$. Suppose that the number of data is n. The likelihood function is given by

$$L(\alpha) = \prod_{i=1}^{n} \frac{1+\alpha x_i}{2}$$

then the log-likelihood is derived as

$$\ell(\alpha) = \sum_{i=1}^{n} \log(1+\alpha x_i) - n\log 2 \tag{3.2}$$

Notice that $-1 \le \alpha x_i \le 1$. $\ell(\alpha)$ is sufficient smooth to implement the Fisher's scoring algorithm. Then we have

$$\ell'(\alpha) = \sum_{i=1}^{n} \frac{x_i}{1+\alpha x_i}$$

$$\ell''(\alpha) = -\sum_{i=1}^{n} \frac{x_i^2}{(1+\alpha x_i)^2} \tag{3.3}$$

and

$$I(\alpha) = -E\big(\ell''(\alpha)\big) = \int_{-1}^{1} \sum_{i=1}^{n} \frac{x^2}{(1+\alpha x)^2} \frac{1+\alpha x}{2} dx = n \int_{-1}^{1} \frac{x^2}{(1+\alpha x)^2} \frac{1+\alpha x}{2} dx$$

$$= \frac{n}{2} \int_{-1}^{1} \frac{x^2}{1+\alpha x} dx = \frac{n}{2\alpha^2} \log \frac{1+\alpha}{1-\alpha} - \frac{n}{\alpha^2}. \tag{3.4}$$

To procced the Newton iteration, it is necessary to evaluate $\ell'(\alpha)$ and $\ell''(\alpha)$. By replacing $\ell''(\alpha)$ to $I(\alpha)$, We obtain the Fisher's scoring algorithm. The evaluating cost of $I(\alpha)$ is smaller than $\ell''(\alpha)$.

4. Two-Step methods

For solving univariate nonlinear equation:

$$\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}, \tag{4.1}$$

we consider some iterative methods based on the Newton mehod[2,3]:

$$x_{k+1} = x_k - g(x_k)/g'(x_k). \tag{4.2}$$

A simple improving scheme is formulated as follow[3-6]:

$$y = x_k - g(x_k)/g'(x_k), \;\; x_{k+1} = y - g(y)/g'(x_k) \tag{4.3}$$

which is a two-step method with 3rd order convergence rate. It necessary to evaluate $g(x)$ twice in (4.3). Another modification is formulated by evaluating $g'(x)$ twice in the two steps. For instance, we have the following two-step iteration:

$$y = x_k - g(x_k)/g'(x_k), \;\; x_{k+1} = y - 2g(x_k)/[g'(x_k) + g'(y)]. \tag{4.4}$$

5. Present methods

Let us present three methods in this section. Those are the two-step method like as (4.3) and (4.4). We present the following two-step methods:

Method A : $\beta = \alpha_k + \ell'(\alpha_k)/I(\alpha_k), \quad \alpha_{k+1} = \beta + \ell'(\beta)/I(\alpha_k)$

Method B: $\beta = \alpha_k + \ell'(\alpha_k)/I(\alpha_k), \quad \alpha_{k+1} = \beta + 2\ell'(\alpha_k)/[I(\alpha_k) + I(\beta)]$

Method C: $\beta = \alpha_k + \ell'(\alpha_k)/I(\alpha_k), \quad \alpha_{k+1} = \beta + 6\ell'(\alpha_k)/[5I(\alpha_k) + I(\beta)]$

6. Numerical experiments and conclusion

Let us examine proposed methods Method A, Method B and Method C. The test problem is the muon decay with the following the data[1] :

X := [.41040018, .91061564, -.61106896, .39736684, .37997637, .34565436, 0.01906680, -.28765977,
-.33169289, .99989810, -.35203164, .10360470, .30573300, .75283842,-.33736278, -.91455101,
-.76222116, .27150040,-0.01257456, .68492778,-.72343908, .45530570, .86249107, .52578673,
.14145264, .76645754, -.65536275, .12497668, .74971197, .53839119].

The starting value is 0.9 and the convergence criteria is $|\alpha_{k+1} - \alpha_k| < 10^{-15}$. The number of iterations for each method becomes

Fisher's scoring: 14, Method A: 8, Method B: 12 and Method C: 8.

From the numerical experiment, Method A and Method C give least number of iterations among them. $I(\beta)$ is evaluated by (3.4) which is involved only the parameter and the number of data. On the other hand, Method A is necessary to evaluate $\ell'(\beta) = \sum_{i=1}^{n} \frac{x_i}{1+\beta x_i}$ which takes the sum about the data in each second step of iterations. Therefore, the computational cost of Method C is less than Method A .

References

1) J.L. Devore and K.N. Berk, Modern mathematical statistics with applications, Duxbury Pr., 2007, ISBN 0534404731.

2) Traub J.F., Iterative Methods for the Solution of Equations. Prentice Hall, New York, 1964.

3) Ostrowski A.M., Solution of Equations and Systems of Equations. Academic Press, New York-London, 1966.

4) Frontini M., Sormani E., Some variant of Newton's method with third-order convergence. Appl. Math. Comput. 140: (2003) 419–426.

5) Shin Min Kanga , Faisal Alib and Arif Rafiqc, Iterative methods for solving scalar equations, J. Nonlinear Sci. Appl. 9 (2016), 1035–1042.

6) S. Weerakon, T. G. I. Fernando, A variant of Newton's method with accelerated third-order convergence. Appl. Math. Lett.,13(2000) 87-93.