日水研報告, (42) : 17—24, 1992.
Bull. Japan Sea Natl. Fish. Res. Inst., (42) : 17—24, 1992.

Mathematical Analysis of Age-Length Key Method for Estimating Age Composition from Length Composition

Tatsuro AKAMINE[1] and Yoshiharu MATSUMIYA[2]

Abstract

In mathematics, the age-length key method is an estimating method for the ratio of each distribution in a mixture of distributions. Although the maximum likelihood method is the principle of this estimation, this model is a nonlinear model with a restrictive condition. At first, HASSELBLAD solved this model by using a new iteration method with a computer. Although his method was regarded as the steepest descent method or EM algorithm, it is proved to be an application of LAGRANGE's indeterminate multiplier method to the iteration method. Nowadays, we have other useful algorithms of optimization, especially MARQUARDT's method, for this estimation.

Key words age-length key, optimization, indeterminate multiplier, EM algorithm, MARQUARDT's method

Introduction

Age-length key is a famous term for fishery population dynamics. This key automatically provides age composition from length composition (SHIMAZU 1980). It is a nonlinear model with a restrictive condition. However, the algorithm for estimating parameters and the existence of the solution for this madel have not yet been sufficiently studied mathematically. HASSELBLAD (1966) solved the madel by employing a computerized iteration method. The algorithm of his method has not been well understood because he mistakenly regarded it as the steepest descent method. ORCHARD and WOODBURY (1972) obtained the same iteration method as HASSELBLAD by using the EM algorithm. Furthermore, KIMURA and CHIKUNI (1987) obtained the same iteration method as FUKUDA and CHIKUNI (unpublished paper) by using the EM algorithm. This method is also identical to HASSELBLAD's method. However, the convergency of the EM algorithm has not been proved (MIYAKAWA 1987). In this paper, we will prove that HASSELBLAD's method is an application of LAGRANGE's indeterminate multiplier method for this nonlinear model which has a restrictive condition. On the other hand, MAKO and MATSUMIYA (1977) developed another

Accepted : November 20, 1991. Contribution A *No.* 474 from Japan Sea National Fisheries Research Institute.

1) Japan Sea National Fisheries Research Institute, Suido-cho, Niigata 951, Japan.
(〒951 新潟市水道町1丁目5939-22 日本海区水産研究所)
2) Faculty of Bioresources, Mie University, Kamihara-cho, Tsu 514, Japan.
(〒514 津市上浜町1515 三重大学生物資源学部)

iteration method for estimating age composition from market size composition. We will also prove that their method is mathematically equal to HASSELBLAD's method. Recently, AKAMINE(1985) successfully applied MARQUARDT's optimization method for this model. We will also explain his method.

Model

1. Simultaneous equations model

The relation of age composition and length composition is defined by the following simultaneous equations:

$$\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} = \begin{pmatrix} f_1(1) & \cdots\cdots & f_n(1) \\ & & \\ f_1(m) & \cdots\cdots & f_n(m) \end{pmatrix} \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} F. \tag{1}$$

Where C_j: length composition (size of length category j),

$f_i(j)$: age-length key of age category i (ratio of length category j in age category i),

P_i: age composition (ratio of age category i to total),

F: total size.

The goal is to solve for P_i with the following restrictive condition:

$$\sum_{i=1}^{n} P_i = 1. \tag{2}$$

Age-length key $f_i(j)$ is regarded as a length frequency distribution in age category i. It is generalized as a probability:

$$\sum_{j=1}^{m} f_i(j) = 1. \tag{3}$$

Combining Equations (1), (2) and (3) results in the following equation:

$$\sum_{j=1}^{m} C_j = F. \tag{4}$$

Equations (2) and (4) indicate that the degrees of freedom for $\boldsymbol{P}$ and $\boldsymbol{C}$ are $(n-1)$ and $(m-1)$, respectively. Therefore, the simultaneous equations in (1) have 3 cases for the existence of the solution. In the first case, when $m < n$, there are so many solutions that it is impractical to solve the madel. In the second case, when $m = n$, the model involves linear simultaneous equations, and there is only one solution. It is easy to obtain this solution (ex. by GAUSS' elimination method). However, this case is involved in the third case, so it can be omitted.

In the third case, when $m > n$, there is no solution for Equations (1). Therefore, the optimum values of $\boldsymbol{P}$ are searched for these equations. The searching method is given by the general optimization. Although DOI(1974) used least squares method for the objective function, the maximum likelihood method is statistically the best method.

2. Maximum likelihood method

From Equations (1), let a mixture of destributions be

$$g(j) = \sum_{i=1}^{n} P_i f_i(j). \tag{5}$$

This is also a probability:

$$\sum_{j=1}^{m} g(j) = 1. \tag{6}$$

Therefore, we can use the maximum likelihood method for estimating parameters P_i. Let the simultaneous probability of the data $\boldsymbol{C}$ be

$$L = \prod_{j=1}^{m} g(j)^{Cj}. \tag{7}$$

This is called the likelihood. The maximum likelihood method makes the parameter values of the maximum L be the best estimators. This method is the principle of parameter estimation for the probability distribution. Because L is too small a positive number to deal with, we define Y as follows:

$$Y = \ln L = \sum_{j-1}^{m} C_j \ln g(j). \tag{8}$$

This large negative number is called the logarithm likelihood.

Although $f_i(j)$ is a discrete model, a continuous model $f_i(x)$ is given by substituting j with x as follows:

$$\int_{\infty}^{\infty} f_i(x)\ \mathrm{d}x = 1. \tag{9}$$

$$g(x) = \sum_{i=1}^{n} P_i f_i(x). \tag{10}$$

$$\int_{\infty}^{\infty} g(x)\ \mathrm{d}x = 1. \tag{11}$$

These equations also satisfy Equations (7) and (8).

In this case, Equation (8) has the restrictive condition (2). Therefore, "conditional optimization" is necessary. HASSELBLAD (1966) solved this problem by using a new iteration method. His model was a special case in which each distribution is a normal distribution as follows:

$$f_i(x) = N(\mu_i,\ \sigma_i). \tag{12}$$

He was able to obtain values not only for each P_i but also for μ_i and σ_i. Therefore, he treated a more general case in "data analysis". Although he called his method the "steepest descent method", it is an application of LAGRANGE's "indeterminate multiplier method" to the iteration method.

Iteration method

1. General theory

The iteration method is mathematically defined as a search for values of θ which satisfy the following equation:

$$h(\theta) = 0. \tag{13}$$

Manipulating Equation (13) leads to

$$\theta = k(\theta). \tag{14}$$

The iteration method is based on this equation as follows:

$$\theta^{\mathrm{new}} = k(\theta^{\mathrm{old}}). \tag{15}$$

Bacause Equation (14) has many variations, this iteration formula does not always converge. Only when the "principle of contracting mapping" is satisfied does this iteration

converge into the "fixed point", which is the solution to Equation (13). For a single parameter, IRI(1981) showed all cases of iteration (15) for convergence or divergence.

In practice, Equation (14) is important. For example, NEWTON's methods provides a useful concrete formula. However, in the case of many parameters NEWTON's method often diverges. MARQUARDT's method, which has a correction factor and is regarded as an expansion of NEWTON's method, is useful for applications involving many parameters.

The maximum point of Equation (8) is the solution to the following simultaneous equations:

$$\frac{\partial Y}{\partial P_i}=0,\ (\ i\ =1\sim n\ ). \tag{16}$$

However, these equations have the restrictive condition of (2). From these equations and the restrictive condition, HASSELBLAD(1966) obtained the follwing iteration formula:

$$P_i^{\text{new}}=\ \left[P_i\sum_j \frac{C_j}{\sum_i P_i\, f_i(j)}\, f_i(j)\right]^{\text{old}}\ /\ F,\ (i=1\sim n-1). \tag{17}$$

This is a very useful iteration method. If $P_i^{\text{old}}\gtrless 0$, then $P_i^{\text{new}}\gtrless 0$. He compared the precision of his method with NEWTON's method, where the initial values were given by his method. For the special case of Equation (12), AKAMINE (1987a) developed a BASIC program, and AKAMINE (1987b) tried to analyse this method. In the next section, we will show that this method is an application of LAGRANGE's indeterminate multiplier method.

2. LAGRANGE's indeterminate multiplier method

In the case of two parameters for linear models, let the objective function be $Y(x,\ y)$ and the restrictive conditon be

$$H(x,\ y)=0. \tag{18}$$

Where x and y are parameters. The objective is to search for the maximum or minimum point of Y with the restrictive condition of (18). Let

$$Q(x,\ y)=Y(x,\ y)-\lambda H(x,\ y). \tag{19}$$

The Solution $(x,\ y,\ \lambda)$ of the simultaneous equations (18) along with

$$\frac{\partial Q}{\partial x}=0,\qquad \frac{\partial Q}{\partial y}=0 \tag{20}$$

gives the object point of $(x,\ y)$. The parameter λ is constant and called the "indeterminate multiplier"

This method is usually applied to linear models. HASSELBLAD's method is regarded as an application of this method to the nonlinear model in (8). In this case , the objective function is Equation (8), and the restrictive condition is

$$H=\sum_{i=1}^{n} P_i-1=0. \tag{21}$$

Where the parameters are $P_i\ (\ i\ =1\sim n\ )$. From Equations (8) and (21),

$$Q=Y-\lambda H. \tag{22}$$

Therefore, from

$$\frac{\partial Q}{\partial P_i}=0, \ (i=1\sim n) \tag{23}$$

we get

$$\lambda=\sum_j C_j \frac{f_i(j)}{g(j)}. \ (i=1\sim n) \tag{24}$$

Because λ is constant, it must be a more simple formula. Multiplying both sides of this equation by P_i leads to

$$P_i\lambda=\sum_j C_j \frac{P_i f_i(j)}{g(j)}. \ (i=1\sim n) \tag{25}$$

Integrating each side of this equation for $i=1\sim n$ leads to

$$\lambda=\Sigma C_j=F \ (\text{constant}). \tag{26}$$

Substituting this value into Equation (25) leads to HASSELBLAD's method (17). For calculating of P_n, he used the following relationship:

$$P_n=1-\sum_{i=1}^{n-1} P_i, \tag{27}$$

instead of the iteration formula (17). Both methods provide the same solution.

3. EM algorithm and another iteration method

In 1974, FUKUDA and CHIKUNI (unpublished paper) invented an algorithm called iterated age-length key (IALK) for estimating age composition from length composition. KIMURA and CHIKUNI (1987) obtained the following iteration formula, which is identical to the IALK algorithm, by using the EM algorithm.

$$P_i^{\text{new}}=\left[\sum_j \frac{C_j}{F} \frac{P_i f_i(j)}{\sum_i P_i f_i(j)}\right]^{\text{old}} \tag{28}$$

This iteration formula is obviously equal to HASSELBLAD's iteration formula (17). On the other hand, ORCHARD and WOODBURY (1972) already obtained formula (17) by using the EM algorithm.

The EM algorithm is an iteration method consisting of an expectation step (E step) and a maximization step (M step). This algorithm was developed for data involving missing values. In E step, missing values are estimated using parameter values. In M step, parameter values are estimated by data involving missing values. The EM algorithm mutually iterates these two steps.

However, the age-length key model (1) has no missing values. Both KIMURA and CHIKUNI (1987) and ORCHARD and WOODBURY (1972) used dummy parameters for missing values. For dummy parameters, the former study employed the size of age category i, and the latter employed the probability beloging age category i in length category j. Finally, they obtained the same formula as HASSELBLAD (1966).

MIYAKAWA (1987) said that the greatest contribution of the EM algorithm is its application to the finite mixture of distributions, which is identical to Equation (5). However, HASSELBLAD's method results from LAGRANGE's indeterminate multiplier method, independent from the EM algorithm shown in this paper. On the other hand,

MIYAKAWA (1987) said that there is no evidence for the convergency of the EM algorithm. It may be proved only by mathematically applying the principle of contraction mapping. Although the EM algorithm only produced HASSELBLAD's method, we have many useful optimization methods for this model; there is no reason to use the EM algorithm.

On the other hand, MAKO and MATSUMIYA (1977) developed another iteration method for estimating age composition from market size composition. MATSUMIYA (1990) suggested that their method is identical to the KIMURA and CHIKUNI method. In this paper, we prove that the MAKO and MATSUMIYA method is essentially equal to HASSELBLAD's method, which is equal to KIMURA and CHIKUNI method.

The MAKO and MATSUMIYA method is an iteration as follows:

$$X_{ij}^{\text{new}} = \left[\frac{X_{ij}}{\sum_j X_{ij}} \sum_j \left(X_{ij} \frac{C_j}{\sum_i X_{ij}}\right)\right]^{\text{old}} \tag{29}$$

Where X_{ij}: estimated size of age category i and length (market size) category j. In this paper, X_{ij} is written as follows:

$$X_{ij} = P_i f_i(j) F. \tag{30}$$

substituing Equation (30) into Equation (29) results in Equation (17). Therefore, the MAKO and MATSUMIYA method is essentially equal to HASSELBLAD's method.

Optimization

AKAMINE (1985) applied MARQUARDT's method to the mixture of normal distributions, the same as HASSELBLAD (1966). In AKAMINE's (1985) "Program 2" and AKAMINE's (1987a) "Appendix A", he used relation (27) to avoid the restrictive condition of (2). However, these BASIC programs contain a small bug, and convergence is a little slow. Althogh the precision of these programs is not a problem in practice, AKAMINE and KATO (1988) corrected the bug.

AKAMINE (1985) used another way to avoid the restrictive condition of (2) in "Program 1". He used the new paramters K_i $(i=1\sim n)$, which are defined by the following equation:

$$P_i = K_i \,/\, \sum_{i=1}^{n} K_i. \tag{31}$$

Although point $\boldsymbol{P}$ converges to a single point in the $(n-1)$-dimensional space, point $\boldsymbol{K}$ converges to a single point on a line in the n-dimensional space. The line is defined by

$$K_i = P_i(\Sigma K_i). \tag{32}$$

Where P_i is constant and calculated in Equation (31).

There are currently many optimization methods for this model. MARQUARDT's method is particularly useful (AKAMINE 1985, 1987a; AKAMINE and KATO 1988). Although HASSELBLAD's method has a large area of convergence, the precision of the solution is not very high. HASSELBLAD (1966) noted this phenomenon after comparing his method to NEWTON's method. AKAMINE (1987a) compared HASSELBLAD's method to MARQUARDT's method and stated that the former needs 200 or more interations for high precision. However, HASSELBLAD's method is a simple algorithm and requires minimal program

memory and time for a single iteration. Therefore, HASSELBLAD's method requires less overall time. It is one of the best practical methods for this model.

The BASIC programs of AKAMINE's(1985) "Program 1" and AKAMINE and KATO (1988), which are MARQUARDT's method, as well as AKAMINE's(1987a) "Appendix B", which is HASSELBLAD's method, are all applicable to this model. However, they are a special model of Equation (12). Therefore, the user must rewrite these programs for the age-length key model; it is not a difficult task.

Conclusion

The age-length key model is a nonlinear model with a restrictive condition. The principle of parameter estimation is the maximum likelihood method. HASSELBLAD's method is applicable to this model. It is an application of LAGRANGE's indeterminate multiplier method, not the steepest descent method or EM algorithm. MARQUARDT's optimization method is also applicable. It requires a technique to avoid the restrictive condition.

Acknowledgments

We are very grateful to the late Mr. F. KATO and Mr. K. ISHIOKA of Nansei National Fisheries Research Institute for their kind advice and help. We are also grateful to Mr. Y. HIYAMA of Japan Sea National Fisheries Research Institute for his critical reading of the manuscript.

References

AKAMINE, T. (1985) Consideration of the BASIC programs to analyse the polymodal frequency distribution into normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.*, (35). 129-159 (In Japanese with English abstract).

AKAMINE, T. (1987a) Comparison of algorithms of several methods for estimating parameters of a mixture of normal distributions. *Bull. Japan Sea Natl. Fish. Res. Inst.*, (37), 259-277.

AKAMINE, T. (1987b) Histrical study for analysis of a length composition and growth curves. *Contribution to the fisheries researches in the Japan Sea block*, (11),23-30 (In Japanese).

AKAMINE, T. and KATO, F. (1988) Estimation of parameters for a mixture of distributions. In *Program issue for population analysis by parsonal computer*. National Research Institute of Fisheries Science, 177-188 (In Japanese).

DOI, T. (1974) Introduction for fishery population dynamics. *Monthly report of nihon suisan sigen hogo kyoukai*, (127), 5-17 (In Japanese).

HASSELBLAD,V. (1966) Estimation of parameters for a mixture of normal distributions. *Technometrics*, **8** (3), 431-444.

IRI, M. (1981) Numerical analysis. Asakura shoten, Tokyo, 1-38 (In Japanese).

KIMURA, D. K. and CHIKUNI, S. (1987) Mixture of empirical distributions: an iterative application of the age-length key. *Biometrics*, (43), 23-35.

MAKO, H. and MATSUMIYA,Y. (1977) Method of estimation of age composition derived from market size composition. *Bull. Seikai Natl. Fish. Res. Inst.*, (50), 1-8 (In Japanese with English abstract).

MATSUMIYA, Y. (1990) Method of estimation of age composition from market size composition: Its mathematical examination. *Nippon Suisan Gakkaishi*, **56** (5), 841.

MIYAKAWA, M. (1987) The EM algorithm and its related problems. *Ouyou toukeigaku*, **16** (1), 1-21 (In Japanese).

ORCHARD, T and WOODBURY, M. A. (1972) A missing information principle: thoery and applications. *Proc. 6th Berkeley Symp. on Math. Statist. Prob.*, (1), 697-715.
SHIMAZU, Y. (1980) A method for estimation of age composition by length composition. *Report of Western Japanese ground fish section of GSK in shouwa 54.*, 36-48 (In Japanese).

体長組成から年齢組成を推定する age-length key 法の数学的解析

赤嶺達郎・松宮義晴

Age-length key 法は数学的には混合分布のパラメータである各分布の比率を推定する方法である．統計学的には最尤法を用いればよいが，非線型モデルであり，しかも制限条件がついているため簡単ではない．この問題は HASSELBLAD によって最初に厳密に解決された．彼の方法は計算機を用いた反復法で，最急降下法あるいは EM アルゴリズムと解釈されてきたが，実際にはラグランジュの未定乗数法の応用であることを証明する．現在では最適化法，特にマルカール法によってより一般的に解くことができる．