

RICHARDS 成長関数のパラメータの推定*

伊藤達夫**

1. はじめに

RICHARDS 成長関数は、わが国では大隅⁴⁾ によって林学界に紹介されて以来、林木の成長現象の記述や解析に広く適用されている。そのパラメータは、計算機の発達によって GAUSS-NEWTON 法に基づく最小二乗法^{5,7,9)} を用いて容易に推定できるようになった。

しかし、林学においてこの関数の適用対象となる樹木や林分の成長データは、生育過程の最終段階までは存在しないことが多く、また、変曲点の現われる以前の初期段階を欠いていることもまれではない。このような条件下で個々のサンプルに対して単純に最小二乗法を適用すると、データに含まれる誤差によってパラメータの推定値が不規則に変動することがある。単にデータの平滑化を目的としているのであればそれでもよいが、パラメータの変化に何らかの法則性を見いだそうとする場合や、成長予測を行なおうとする場合には、困惑することになる。このような問題は、RICHARDS 成長関数の柔軟性の代償であるパラメータの不安定性から生じる。ここで注意しなければならないのは、当然のことであるが、残差平方和はパラメータ空間内で連続的に変化しているので、最小点の回りにはその値があまり変わらない部分が広がっているということである。つまり、最小二乗推定値はそれなりに意味を持つものではあるが、偶然的に定まった一つの点であり絶対的なものではないということである。このことを考慮すれば、RICHARDS 関数の応用の目的によっては、個々のサンプルに対する最小二乗推定値にこだわることは合理的ではないといえる。

それでは、パラメータの推定値が不安定になることを防ぎつつ、この関数を成長解析や成長モデルの構築に活用するにはどのようにしたらよいか。一つの方法は、データから推定するパラメータの数をなるべく減らすことである。いくつかのパラメータの値を固定することができれば、残りのパラメータの推定値は安定する。あるパラメータの平均的な値が多くあてはめ結果からわかっている場合には、それを固定値として用いることができるし、成長要素間に仮定される関数関係からパラメータの値を決定できることもある。このパラメータの値を固定するという方法は、成長解析に有効である。例えば、十分成育した段階までのデータがないために成長曲線の漸近値を示すパラメータの推定値が極端に大きくなった場合には、残りのパラメータの推定値もそれに

*Estimating parameters of the RICHARDS growth function

**Tatsuo Ito, Faculty of Agr., Kyoto Pref. Univ., Kyoto 606 京都府立大学農学部

引かれて片寄っているので、それらを他の安定したデータからの推定値と比較することができない。ここで、どれか一つのパラメータの値をこの安定したデータからのものと共通であると仮定することができれば、それを固定値として用いて推定をやり直して、残りのパラメータの推定値と比較することができる。再度のあてはめによる残差平方和の増大が小さければ、このやり方は十分正当化されるだろう。

パラメータの推定値を安定させるより合理的な方法は、値が共通なパラメータを持つと仮定される複数のサンプルを一括して同時に推定を行なうことである。この場合、個々のサンプルは共通なパラメータによって相互に関連付けられるので、最小二乗法は各サンプルに対する残差平方和の総和を最小化するように働く。例えば、あるパラメータの平均的な値を求めたい場合には、個々のサンプルに対する最小二乗推定値を算術平均するよりも、そのパラメータの値が全サンプルに共通であると仮定してこの方法で推定する方が合理的である。また、成長解析に際しては、共通な値を持つパラメータがあると仮定することができれば、サンプル間の成長の差異を残ったパラメータに集約して表わすことができる。

値が共通なパラメータを持つ複数のサンプルを一括して推定を行なう方法は、このように RICHARDS 成長関数の応用に有用であるが、まだ一般的には適用されていない。本報告ではこの方法の詳細について述べる。以下、まず、推定方法の基礎となる GAUSS-NEWTON 法の概要を、RICHARDS 成長関数を単一のサンプルに対してあてはめる場合について説明する。次に、複数のサンプルの場合についても同様に GAUSS-NEWTON 法が適用できることを示し、計算機プログラム作成上の要点を解説する。

2. パラメータ推定の基礎

RICHARDS 成長関数は次の式で表される⁶⁾。

$$W = A(1 - be^{-kt})^{1/(1-m)} \quad (1)$$

ただし、 W は時間 t における成長要素の大きさであり、 A , b , k , m はパラメータである。この関数の解析的な性質や各パラメータの意味については他の文献^{1,5,6)}を参照されたい。積分定数であるパラメータ b はこのままでは扱いにくいので、初期条件を与えると式 (1) は次のように変形できる。

$$W = A(1 - e^{-k(t-t_0)})^{1/(1-m)} \quad (2)$$

$$W = A[1 - \{1 - (W_0/A)^{1-m}\}e^{-kt}]^{1/(1-m)} \quad (3)$$

ただし、式 (2) においては、 $m < 1$ であり $t = t_0$ のとき $W = 0$ である。また、式 (3) においては、 $m > 1$ であり $t = 0$ のとき $W = W_0$ である。

通常、パラメータ t_0 や W_0 の値は、データから直接推定せずに、他のデータから推定したり経験的に決定した値を固定値として与えることが望ましい。そうすることによって、他のパラメータの値が不安定になることを少しでも防ぐことができる。例えば、胸高直径の成長データに式 (2) を

あてはめる場合には、樹高成長のデータから推定した樹高が胸高に達する年齢に t_0 の値を固定することができる。ただし、サンプルの大きさが小さい場合には、これらのパラメータを不用意な値に固定すると、他のパラメータの推定値が大きく片寄ることがあるので注意が必要である。

以下、式(2)の場合について GAUSS-NEWTON 法によるパラメータの推定方法を説明するが、式(1)、(3)の場合も同様である。

n 組のデータ (W_j, t_j) ($j=1, 2, \dots, n$) を含む 1 個のサンプルから

$$W_j = A(1 - e^{-k(t_j - t_0)})^{1/(1-m)} + \varepsilon_j \quad (4)$$

ただし、 $\varepsilon_j \sim N(0, \sigma^2)$

と仮定して、4つのパラメータ

$$x = \begin{pmatrix} A \\ m \\ k \\ t_0 \end{pmatrix} \quad (5)$$

を推定することを考える。

式(4)の仮定より、この n 組のデータに対する尤度は

$$\begin{aligned} L(A, m, k, t_0, \sigma^2) \\ = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp \left[-\frac{1}{2\sigma^2} \sum_{j=1}^n \{ W_j - A(1 - e^{-k(t_j - t_0)})^{1/(1-m)} \}^2 \right] \end{aligned} \quad (6)$$

対数尤度は

$$\begin{aligned} \log L(A, m, k, t_0, \sigma^2) \\ = l(A, m, k, t_0, \sigma^2) \\ = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{j=1}^n \{ W_j - A(1 - e^{-k(t_j - t_0)})^{1/(1-m)} \}^2 \end{aligned} \quad (7)$$

となる。式(7)の値を最大にするパラメータを求めるのが最尤法であり、そのためには式中の残差平方和

$$\sum_{j=1}^n \{ W_j - A(1 - e^{-k(t_j - t_0)})^{1/(1-m)} \}^2$$

を最小にしなければならない。したがって、この場合、最尤法は最小二乗法となる。

ここで、残差ベクトル $f(x)$ と最小化すべき目的関数 $S(x)$ を以下のように定義する。

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} = \begin{pmatrix} W_1 - A(1 - e^{-k(t_1 - t_0)})^{1/(1-m)} \\ W_2 - A(1 - e^{-k(t_2 - t_0)})^{1/(1-m)} \\ \vdots \\ W_n - A(1 - e^{-k(t_n - t_0)})^{1/(1-m)} \end{pmatrix} \quad (8)$$

$$S(x) = \frac{1}{2} \sum f_j^2(x) \quad (9)$$

束領域が非常に小さい。

式(15)において、 $\sum f_j(x^k) \nabla^2 f_j(x^k)$ の項を無視し、ヘッセ行列 $\nabla^2 S(x^k)$ を $J^T(x^k)J(x^k)$ で近似することにすれば、以下の反復公式を得る。

$$x^{k+1} = x^k - \{J^T(x^k)J(x^k)\}^{-1} J^T(x^k) f(x^k) \quad (16)$$

これを GAUSS-NEWTON 法という。解 \bar{x} における残差 $f(\bar{x})$ が小さい場合や、 $f(x)$ が線形に近い場合、あるいは、 $f_j(x) \nabla^2 f_j(x)$ が互いに打ち消しあって $\sum f_j(x) \nabla^2 f_j(x)$ が小さくなる場合は、式(16)に基づく方法は効率的である⁸⁾。

式(16)において、パラメータの微調整量のベクトル $x^{k+1} - x^k$ を δ^k とおけば、

$$\delta^k J^T(x^k) J(x^k) = -J^T(x^k) f(x^k) \quad (17)$$

適当な初期値 x^0 から出発して、反復計算の各回においてこの連立方程式を解いて δ^k を求め、 x^k を補正する。残差平方和の減少量が十分小さくなり、また δ^k の絶対値も十分小さくなったら計算を打ち切る。反復計算を停止する具体的な基準については、今野・山下²⁾、中川・小柳³⁾ を参照されたい。

$J^T(x^k)J(x^k)$ は、 4×4 行列であり、その要素は、 $\partial f_j / \partial A = \alpha_j$ 、 $\partial f_j / \partial m = \beta_j$ 、 $\partial f_j / \partial k = \gamma_j$ 、 $\partial f_j / \partial t_0 = \zeta_j$ とすれば、

$$J^T(x^k)J(x^k) = \begin{pmatrix} \sum \alpha_j^2 & \sum \alpha_j \beta_j & \sum \alpha_j \gamma_j & \sum \alpha_j \zeta_j \\ \sum \alpha_j \beta_j & \sum \beta_j^2 & \sum \beta_j \gamma_j & \sum \beta_j \zeta_j \\ \sum \alpha_j \gamma_j & \sum \beta_j \gamma_j & \sum \gamma_j^2 & \sum \gamma_j \zeta_j \\ \sum \alpha_j \zeta_j & \sum \beta_j \zeta_j & \sum \gamma_j \zeta_j & \sum \zeta_j^2 \end{pmatrix} \quad (18)$$

縦ベクトル $J^T(x^k)f(x^k)$ の要素は

$$J^T(x^k)f(x^k) = \begin{pmatrix} \sum \alpha_j f_j(x^k) \\ \sum \beta_j f_j(x^k) \\ \sum \gamma_j f_j(x^k) \\ \sum \zeta_j f_j(x^k) \end{pmatrix} \quad (19)$$

である。式(18)、(19)の行列は、ヤコビ行列と残差ベクトルを作成してからそれらの内積として得るのではなく、各要素を直接計算して得た方が効率がよい。これらの行列の大きさは、値を固定するパラメータがある場合はその数に応じて小さくなる。

式(16)による反復計算の開始に必要なパラメータの初期値は、すべてを試行錯誤的に与えることもできるが、相互のバランスをとるのが困難であり発散することが多く効率的でない。収束の安定化のためには、パラメータ m の値を固定して残りのパラメータの推定をその固定値を変化させながら繰り返し、残差平方和の最小になる区間がわかったところで m を含む推定を行なうという方法⁷⁾がある。この方法は確実性があるがやや煩雑であるので、ここで推奨する以下の方法がうまくいかなかった場合に用いるとよい。パラメータの中で直感的に値を定めることが最も困難であるのは k であるので、その初期値だけを他のパラメータの初期値を用いて次の式から求める

ようにすると発散することが少なくなる。

$$k_0 = -\frac{\sum \log\{1 - W_j/A^0\}^{(1-m_0)}\{t_j - t_0\}}{\sum (t_j - t_0)^2} \quad (20)$$

ただし、 $W_j \geq A^0$ の成り立つデータは、この計算から除外しなければならない。

収束させることが不可能であるかどうかの判断は困難であるが、多くの場合、パラメータ A の固定値を変化させながらあてはめを繰り返し、残差平方和の動きを見ることによって見当をつけることができる。任意のパラメータの値を固定してあてはめを行なうことができるプログラムを作っておくとこのような場合に役立つ。

3. 複数のサンプルからの制約付きの推定

k 個のサンプルがあり、それぞれが n_i ($i=1, 2, \dots, k$) 組のデータ (w_{ij}, t_{ij}) ($j=1, 2, \dots, n_i$) を含んでいるとする。このデータについて、次のように仮定する。

$$W_{ij} = A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)} + \varepsilon_{ij} \quad (21)$$

ただし、 $\varepsilon_{ij} \sim S(0, \sigma^2_i)$

それぞれのサンプルに対する RICHARDS 成長関数の 4 つのパラメータを

$$x_i = \begin{pmatrix} A_i \\ m_i \\ k_i \\ t_{0i} \end{pmatrix} \quad (22)$$

とし、その推定を考える。

対数尤度は次のようになる。

$$\begin{aligned} & l(x_1, x_2, \dots, x_k, \sigma^2_1, \sigma^2_2, \dots, \sigma^2_k) \\ &= -\frac{1}{2} \sum_{i=1}^k n_i \log 2\pi - \frac{1}{2} \sum_{i=1}^k n_i \log \sigma^2_i \\ & \quad - \frac{1}{2} \sum_{i=1}^k \left[\frac{1}{\sigma^2_i} \sum_{j=1}^{n_i} \{W_{ij} - A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)}\}^2 \right] \end{aligned} \quad (23)$$

分散がサンプルごとに異なり、しかもサンプル間で共通な値をもつと仮定されるパラメータがある場合には、この対数尤度の最大化に最小二乗法は適用できず、式 (23) を直接目的関数とする非線形最適化計算を行わなければならない。それには困難な数値計算が伴うので、ここでは分散は全サンプルについて一定とみなすことができるということを前提条件とする。この条件により、ここでも最尤法は最小二乗法となる。残差ベクトル $f(\theta)$ と最小化すべき目的関数 $S(\theta)$ は次のように定義される。

$$\begin{pmatrix} f_{11}(x_1) & W_{11} - A_1(1 - e^{-k_1(t_{11} - t_{01})})^{1/(1-m_1)} \\ f_{12}(x_1) & W_{12} - A_1(1 - e^{-k_1(t_{12} - t_{01})})^{1/(1-m_1)} \\ \vdots & \vdots \end{pmatrix}$$

$$f(\theta) = \begin{pmatrix} f_{1n_1}(x_1) & W_{1n_1} - A_1(1 - e^{-k_1(t_{1n_1} - t_{01})})^{1/(1-m_1)} \\ f_{21}(x_2) & W_{21} - A_2(1 - e^{-k_2(t_{21} - t_{02})})^{1/(1-m_2)} \\ f_{22}(x_2) & W_{22} = A_2(1 - e^{-k_2(t_{22} - t_{02})})^{1/(1-m_2)} \\ \vdots & \vdots \\ f_{2n_2}(x_2) & W_{2n_2} - A_2(1 - e^{-k_2(t_{2n_2} - t_{02})})^{1/(1-m_2)} \\ \vdots & \vdots \\ f_{k1}(x_k) & W_{k1} - A_k(1 - e^{-k_k(t_{k1} - t_{0k})})^{1/(1-m_k)} \\ f_{k2}(x_k) & W_{k2} - A_k(1 - e^{-k_k(t_{k2} - t_{0k})})^{1/(1-m_k)} \\ \vdots & \vdots \\ f_{kn_k}(x_k) & W_{kn_k} - A_k(1 - e^{-k_k(t_{kn_k} - t_{0k})})^{1/(1-m_k)} \end{pmatrix} \quad (24)$$

$$\begin{aligned} S(\theta) &= \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^{n_i} \{W_{ij} - A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)}\}^2 \\ &= \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^{n_i} f_{ij}^2(x_i) \end{aligned} \quad (25)$$

ここで、 θ は推定するパラメータのベクトルで、その要素数は、サンプル間で共通な値を持つパラメータや値の固定されるパラメータがある場合には、 $k \times 4$ より小さくなる。例えば、3つのサンプルがあり、それらについて、 $m_1 = m_2 = m_3 = m_{(1, 2, 3)}$ 、 $k_2 = k_3 = k_{(2, 3)}$ と仮定され、 t_0 の値がすべてのサンプルについて一定値に固定される場合には、 $\theta = (A_1, A_2, A_3, m_{(1, 2, 3)}, k_1, k_{(2, 3)})^T$ となり、その要素数は6である。この場合、ヤコビ行列 $J(\theta)$ は $(n_1 + n_2 + n_3)$ 行6列であり、その内容は以下ようになる。

$$J(\theta) = \begin{pmatrix} \frac{\partial f_{11}}{\partial A_1} & 0 & 0 & \frac{\partial f_{11}}{\partial m_{(1, 2, 3)}} & \frac{\partial f_{11}}{\partial k_1} & 0 \\ \frac{\partial f_{12}}{\partial A_1} & 0 & 0 & \frac{\partial f_{12}}{\partial m_{(1, 2, 3)}} & \frac{\partial f_{12}}{\partial k_1} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_{1n_1}}{\partial A_1} & 0 & 0 & \frac{\partial f_{1n_1}}{\partial m_{(1, 2, 3)}} & \frac{\partial f_{1n_1}}{\partial k_1} & 0 \\ 0 & \frac{\partial f_{21}}{\partial A_2} & 0 & \frac{\partial f_{21}}{\partial m_{(1, 2, 3)}} & 0 & \frac{\partial f_{21}}{\partial k_{(2, 3)}} \\ 0 & \frac{\partial f_{22}}{\partial A_2} & 0 & \frac{\partial f_{22}}{\partial m_{(1, 2, 3)}} & 0 & \frac{\partial f_{22}}{\partial k_{(2, 3)}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \frac{\partial f_{2n_2}}{\partial A_2} & 0 & \frac{\partial f_{2n_2}}{\partial m_{(1, 2, 3)}} & 0 & \frac{\partial f_{2n_2}}{\partial k_{(2, 3)}} \\ 0 & 0 & \frac{\partial f_{31}}{\partial A_3} & \frac{\partial f_{31}}{\partial m_{(1, 2, 3)}} & 0 & \frac{\partial f_{31}}{\partial k_{(2, 3)}} \end{pmatrix} \quad (26)$$

$$\begin{pmatrix} 0 & , & 0 & , & \frac{\partial f_{32}}{\partial A_3} & , & \frac{\partial f_{32}}{\partial m_{(1, 2, 3)}} & , & 0 & , & \frac{\partial f_{32}}{\partial k_{(2, 3)}} \\ \vdots & & \vdots \\ 0 & , & 0 & , & \frac{\partial f_{3n_3}}{\partial A_3} & , & \frac{\partial f_{3n_3}}{\partial m_{(1, 2, 3)}} & , & 0 & , & \frac{\partial f_{3n_3}}{\partial k_{(2, 3)}} \end{pmatrix}$$

サンプルの種類が異なる (例えば、樹高と胸高断面積) ために分散を一定とみなすことができない場合には、何らかの方法で残差をスケールリングする。それには以下のような方法が考えられる。

1) 百分率残差を用いる

$$f_{ij}(x_i) = 1 - A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)} / W_{ij} \quad (27)$$

2) 各サンプルごとのデータの平均値 \bar{W}_i で割る

$$f_{ij}(x_i) = \{ W_{ij} - A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)} \} / \bar{W}_i \quad (28)$$

3) 制約なしのあてはめによって推定された分散 σ^2_{*i} を用いる

$$f_{ij}(x_i) = \{ W_{ij} - A_i(1 - e^{-k_i(t_{ij} - t_{0i})})^{1/(1-m_i)} \} / (\sigma^2_{*i})^{0.5} \quad (29)$$

これらの方法によって残差ベクトルをスケールリングする場合には、ヤコビ行列の要素の修正が必要になる。

式 (26) の例からもわかるように、複数のサンプルに対する制約付きのあてはめの場合、ヤコビ行列の各列はパラメータベクトル θ の各要素に対応し、その列のパラメータに無関係なサンプルに対応する行については要素が 0 となる。ヤコビ行列と残差ベクトルを作ることができれば、式 (16) と同じ反復公式が適用できる。しかし、サンプル数が増えると、ヤコビ行列を作成することは計算機の記憶容量の点から困難になる。例えば、50 本の樹幹解析木のそれぞれについて 10 個の樹高成長データがあり、これらのデータからパラメータ A の値は木によって異なるが残りのパラメータの値はすべての木に共通であり t_0 の値は 0 であると仮定して推定を行なう場合を考えると、ヤコビ行列の大きさは $(50 \times 10) \times (50 + 1 + 1)$ となり、8 バイト実数で 208 キロバイトの記憶領域を占める。各行で 0 でない要素が入っているのはこの場合常に 3 列だけであることに着目すれば、この行列を 500×3 に圧縮して記憶することができる。これは効率的であるように思えるが、後で内積をとるときに復元に手間をかけなければならないことを考えると、むしろ行列 $J^T(\theta)J(\theta)$ と $J^T(\theta)f(\theta)$ をデータから直接作成した方がよい。これらの行列は、複数のサンプルに共通なパラメータがある場合には内容が複数で、式 (18), (19) のような一般的定義を与えることはできない。しかし、パラメータに関する制約条件を適切なデータ構造で表わすことができれば、汎用性のある計算機プログラムを書くことはそれほど困難ではない。

サンプル数が増えるとすべてのパラメータの初期値を直感的に定めることは難しくなるので、 k の初期値は先に述べたように計算によって与えた方がよい。サンプルが複数の場合でも式 (20) をそのまま使うことができる。ただし、複数のサンプルに共通な k の初期値を求める際には、式 (20) 中のパラメータ A や m の値はサンプルごとに異なる可能性がある。

4. おわりに

RICHARDS 成長関数の応用は、現在、多くのあてはめ結果の帰納的分析の段階から、パラメータの値に関して何らかの仮定・制約を設けた成長モデル構築の段階へと進んで来ている。以上に報告した複数のサンプルからのパラメータの推定方法は、このようなモデルのパラメータの決定に役立つ。モデルにおける制約条件に応じてパラメータの推定を行なうという考え方は、RICHARDS 成長関数を有効に応用していく上で重要である。

引用文献

- 1) FLETCHER, R. I. : A general solution for the complete Richards function. *Math. Biosci.* 27 : 349~360, 1975
- 2) 今野 浩・山下 浩 : 非線形計画法. 345pp, 日科技連, 東京, 1978
- 3) 中川 徹・小柳義夫 : 最小二乗法による実験データ解析. 206pp, 東京大学出版会, 東京, 1982
- 4) 大隈眞一 : RICHARDS の生長函数とその林木生長への応用. 87 回目林論 : 111~112, 1976
- 5) 大隈眞一・石川善朗 : 樹木の生長解析に対する RICHARDS 生長関数の適用性について. 京府大学術報告・農学 35 : 49~76, 1983
- 6) RICHARDS, F. J. : A flexible growth function for empirical use. *J. Exp. Bot.* 10 : 290~300, 1959
- 7) 白石則彦 : Basic による Richards 関数あてはめのプログラム. 林業統計研究会誌 8 : 50~57, 1983
- 8) 田辺国土 : 非線形最小二乗法のアルゴリズム. 応用統計学 9 : 119~140, 1981
- 9) 吉田成章 : 生長曲線の検討. 日林誌 61 : 321~329, 1979