mwp-326

menl - 非線形混合効果モデル　【 評価版 】

menl コマンドは固定効果や変量効果の一部、またはすべてが非線形な形で表現されるような混合効果モデルをフィットさせます。ただし全般的な誤差の分布は正規分布であることが仮定されます。また変量効果、及び最下位グループ内における不均一分散/相関のモデル化を支援するために異なる共分散構造が用意されています。

1. 非線形混合効果モデル

非線形混合効果 (NLME: nonlinear mixed-effects) モデルというのは、固定効果や変量効果の一部、またはすべてが非線形な形で組み入れられたモデルのことを言います。それを線形混合効果 (LME: linear mixed-effects) モデル（`[ME]` **mixed** (*mwp-018*) 参照）の一般化ととらえる場合には、変量効果が所与としたときのアウトカムの条件付き平均が係数値と変量効果の非線形関数として表現されると考えることになります。一方、それを独立なデータに対する非線形回帰モデル（`[R]` **nl** (*mwp-178*) 参照）の拡張ととらえた場合には、係数中に変量効果 — 階層のレベル間で異なる値を持ち、それによってレベル内における相関を惹起する — を許容したモデルを想定することになります。

それではなぜ NLME モデルを使うのでしょうか？高次多項式 LME モデルや一般化線形混合効果 (GLME: generalized linear mixed-effects) モデルを代りに使うことでは具合が悪いのでしょうか？

原理的には、滑らかな非線形関数であればどれも高次多項式によって近似が可能です。このため LME 多項式モデルを用い、十分正確な近似が得られるまで多項式の次数を高くしてやれば良いと主張する人もいるかも知れません。しかしこのアプローチには 3 つの問題があります。第 1 点は NLME モデル中のパラメータが通常、半減期等の物理的な概念と結び付いているという点です。これに対し LME 多項式モデルの場合にはそうは言えません。第 2 に NLME モデルの方が通常、対応する LME 多項式モデルに比べてパラメータ数が少ないという点があげられます。すなわちデータの集約がより効率的に行える形となります。第 3 の問題点は NLME モデルの方が観測値の範囲を超えたデータに対してより良好な予測をもたらす傾向にあるという点です。

GLME モデル（`[ME]` **meglm** (*mwp-219*) 参照）も非線形なモデルです。しかし条件付きの平均応答は固定効果と変量効果の双方を含む線形予測子の非線形関数であるとは言えても、あくまでその非線形性は間接的なものでしかありません。そうは言っても GLME 用の推定法と NLME 用の推定法の多くは類似性の高いものとなっています。

NLME モデルは母集団薬物動態学 (population pharmacokinetics)、生物検定 (bioassays) 等の分野で使用されてきました。関係する文献には次のようなものがあります。

- Davidian and Giltinan (1995)
- Vonesh and Chinchilli (1997)
- Demidenko (2013)
- Pinheiro and Bates (2000)
- Davidian and Giltinan (2003)

今、関心対象の母集団から抽出された M 人（個）の観察対象 (subjects) からなる標本について考えることにします。ただし観察対象 j に対しては時点 $t_{1j},\dots,t_{n_j j}$ における合計 n_j 個の計測値 $y_{1j},\dots,y_{n_j j}$ が存在するものとします。この場合、“観察対象”というのは複数の相関を持った観測データ (observations) からなる実験単位や個人、パネル、クラスタを意味するわけです。基本的な非線形の 2 レベルモデルは次のように記述されます [*1]。

$$y_{ij} = \mu(\mathbf{x}'_{ij}, \boldsymbol{\beta}, \mathbf{u}_j) + \epsilon_{ij} \qquad i = 1,\dots,n_j; j = 1,\dots,M \tag{1}$$

ただし $\mu(\cdot)$ は $p\times 1$ ベクトルの固定効果 $\boldsymbol{\beta}$、$q\times 1$ ベクトルの変量効果 $\mathbf{u}_j$ — 平均が $\mathbf{0}$、分散共分散行列を $\boldsymbol{\Sigma}$ とする多変量正規分布に従うものとする —、及び共変量ベクトル $\mathbf{x}_{ij}$ — subject 内の共変量 $\mathbf{x}^w_{ij}$ と subject 間の共変量 $\mathbf{x}^b_j$ の双方を含む — に依存する実値の関数とします。また $n_j\times 1$ ベクトルである誤差 $\boldsymbol{\epsilon}_j = (\epsilon_{1j},\dots,\epsilon_{n_j j})$ は平均が $\mathbf{0}$、分散共分散行列を $\sigma^2\boldsymbol{\Lambda}_j$ とする多変量正規分布に従うものとします。ただし σ^2 は $\boldsymbol{\Lambda}_j$ によって群内誤差分散 σ^2_ϵ、またはスケールパラメータ σ^2 のいずれかを意味するものとします。

NLME モデルのパラメータは科学的な意味合いを持つものが多く、研究の課題もそれらに基づき構成されます。この点を配慮しつつ (1) 式をそれと等価な 2 段階階層モデルに書き換えると次のようになります。

$$\begin{aligned} \text{Stage 1: Individual-level model } & y_{ij} = m(\mathbf{x}^w_{ij}, \boldsymbol{\phi}_j) + \epsilon_{ij} \quad && i = 1,\dots,n_j \\ \text{Stage 2: Group-level model } & \boldsymbol{\phi}_j = \boldsymbol{d}(\mathbf{x}^b_j, \boldsymbol{\beta}, \mathbf{u}_j) && j = 1,\dots,M \end{aligned} \tag{2}$$

Stage 1 においては subject 内の挙動を記述する関数 $m(\cdot)$ を用いて応答をモデル化します。この関数は物理的な解釈が可能な subject 固有のパラメータ $\boldsymbol{\phi}_j$、及び subject 内の共変量ベクトル $\mathbf{x}^w_{ij}$ に依存する形となっています。Stage 2 においては subject 間の挙動をモデル化するために、すなわち $\boldsymbol{\phi}_j$ 自身とそれが subject 間でどう変化するかをモデル化するために、既知のベクトル値関数 $\boldsymbol{d}(\cdot)$ を使用します。この $\boldsymbol{d}(\cdot)$ 関数は subject 間共変量ベクトル $\mathbf{x}^b_j$ と変量効果に依存したものとして規定されます。すなわち Stage 1 ではそれぞれの subject ごとに共通の関数形を規定し、次に Stage 2 においてはパラメータの一部が subject 間でランダムに変動することを許容しようという考え方に基づくものと言えます。

[*1] 1 レベルの NLME は独立なデータに対する単なる非線形回帰モデルに他なりません。

(1) 式と (2) 式の意味をつかむために、ここでは大豆に関するデータを例にとって見て行くことにします (Davidian and Giltinan (1995))。具体的には大豆の葉の平均重量が y_{ij} として記録されてわけですが、この場合、j は区画 (plot) に対する、i は時間（植えてからの日数）に対する suffix である点に注意してください。モデル式を (1) の形で表現すると次のようになります。

$$\begin{aligned} y_{ij} &= \mu(\mathbf{x}'_{ij}, \boldsymbol{\beta}, \mathbf{u}_j) + \epsilon_{ij} \\ &= \frac{\beta_1 + u_{1j}}{1 + \exp[-\{t_{ij} - (\beta_2 + u_{2j})\}/(\beta_3 + u_{3j})]} + \epsilon_{ij} \end{aligned}$$

この例の場合、$\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)', \mathbf{u}_j = (u_{1j}, u_{2j}, u_{3j})'$ であり、$\mathbf{x}_{ij}$ には単に t_{ij} が対応することになります。

一方、モデル式を (2) の形で表現すると次のようになります。

$$\begin{aligned} \text{Stage 1: } y_{ij} &= m(\mathbf{x}^w_{ij}, \boldsymbol{\phi}_j) + \epsilon_{ij} \\ &= \frac{\phi_{1j}}{1 + \exp\{-(t_{ij} - \phi_{2j})/\phi_{3j}\}} + \epsilon_{ij} \\ \text{Stage 2: } \phi_{1j} &= \beta_1 + u_{1j} \\ \phi_{2j} &= \beta_2 + u_{2j} \\ \phi_{3j} &= \beta_3 + u_{3j} \end{aligned}$$

ただし $\mathbf{x}^w_{ij} = t_{ij}, \boldsymbol{\phi}_j = (\phi_{1j}, \phi_{2j}, \phi_{3j})' = \boldsymbol{d}(\mathbf{x}^b_j, \boldsymbol{\beta}, \mathbf{u}_j) = \boldsymbol{\beta} + \mathbf{u}_j$ という対応関係があります。

(2) 式の優位性は解釈のしやすさにあります。なぜなら $\boldsymbol{\phi}_j$ は曲線 (trajectory) を特徴付けるパラメータであると言えるからです。例えば ϕ_{1j} は $t_{ij} \to \infty$ としたときの区画 j における漸近的な葉の重量平均値、ϕ_{2j} は ϕ_{1j} の 1/2 の値に至るまでの時間と考えることができます [*2]。`menl` は双方の表現形式に対応しています。

評価版では割愛しています。

2. 変量効果置換可能数式

`menl` によってフィットすべき非線形モデルは変量効果置換可能数式 (random-effects substitutable expression) — 変量効果を含む置換可能数式 — を用いて定義されます。その例としては `exp({b}+{U[id]})`, `{b1}/({b2}+{b3}*x+{U[id]})`, `({b1}+{U1[id]})/(1+{b2}*x+{c.x#U2[id]})` などがあげられます。

2.1 置換可能数式

評価版では割愛しています。

2.2 線形結合

評価版では割愛しています。

2.3 線形形式

評価版では割愛しています。

[*2] $t_{ij} = \phi_{2j}$ とすると $E(y_{ij}) = \phi_{1j}/2$ となる点に注意。

2.4 変量効果

評価版では割愛しています。

2.5 多階層の指定様式

評価版では割愛しています。

2.6 まとめ

評価版では割愛しています。

3. 初期値の設定

評価版では割愛しています。

4. 2 階層モデル

本セクションでは menl による 2 階層モデルのフィットについてその用例を紹介しますが、より基本的な用例については [ME] **me** (*mwp-204*) のセクション 4.5 にも記載されているのでそちらも参照ください。

ここでは Example データセット unicorn.dta を使用します。

```
. use http://www.stata-press.com/data/r15/unicorn.dta *3
(Brain shrinkage of unicorns in the land of Zootopia)
```

このデータセット中には雄と雌の新生一角獣 (unicorns) 各々 30 頭の脳の重量が変数 weight として記録されています。計測は出生後 2 カ月ごと、2 年間にわたって行われているため、それぞれの個体につき 13 個の観測データが存在することになります。時間変数は time ですが単位は年であるため、$0, 1/6, 2/6, \ldots, 2$ という値が小数値で記録されています。一角獣の脳の萎縮 (brain shrinkage) に関するモデルは過去の調査に基づき次のように想定されています。

$$\texttt{weight}_{ij} = \beta_1 + (\beta_2 - \beta_1)\exp(-\beta_3 \texttt{time}_{ij}) + \epsilon_{ij} \qquad i = 1, 2, \ldots, 13; j = 1, 2, \ldots, 60$$

ただし i は時間に関する suffix であり、j は 60 頭の個体に対する suffix である点に注意してください。

パラメータ β_1 は $\texttt{time}_{ij} \to \infty$ としたときの平均的な脳の重量を表します。これに対しパラメータ β_2 は出生時点 ($\texttt{time}_{ij} = 0$) における脳重量の平均値を、β_3 は脳重量の平均値が漸近的重量 β_1 に近づく速度（減衰速度）を規定するスケールパラメータを意味します。このモデルには変量効果が含まれていないので nl コマンド（[R] **nl** (*mwp-178*) 参照）によるフィットが可能です。

ここでは最初に漸近パラメータ β_1 が個体によって変動するモデルを想定することにします。すなわち β_1 を

*3 メニュー操作：File ▷ Example Datasets ▷ Stata 15 manual datasets と操作、Multilevel Mixed-Effects Reference Manual [ME] **の** menl の項よりダウンロードする。

$\beta_1 + u_{0j}$ で置換えたモデル式

$$\texttt{weight}_{ij} = \beta_1 + u_{0j} + (\beta_2 - \beta_1 - u_{0j})\exp(-\beta_3\texttt{time}_{ij}) + \epsilon_{ij} \tag{3}$$

を考察対象とするわけです。この場合、$\beta_1, \beta_2, \beta_3$ は固定効果パラメータであり、u_{0j} は変数 id によって特定される一角獣個体レベルのランダム切片です。ただし u_{0j} は平均が 0、分散が σ_u^2 の正規分布に従うものとします。

(3) 式を等価な 2 段階モデルに書き換えると次のようになります。

$$\texttt{weight}_{ij} = \phi_{1j} + (\phi_{2j} - \phi_{1j})\exp(-\phi_{3j}\texttt{time}_{ij}) + \epsilon_{ij} \tag{4}$$

$$\left.\begin{aligned} \phi_{1j} &= \beta_1 + u_{0j} \\ \phi_{2j} &= \beta_2 \\ \phi_{3j} &= \beta_3 \end{aligned}\right\} \tag{5}$$

この場合、パラメータ $\phi_{1j}, \phi_{2j}, \phi_{3j}$ は j 番目の一角獣の挙動を記述することになります。例えば ϕ_{1j} は $\texttt{time}_{ij} \to \infty$ としたときの j 番目の個体の脳重量を表すわけです。

▷ Example 1: **単純な 2 階層モデル**

最初に (3) 式で表現される単一方程式モデルを menl でフィットさせます。

- Statistics ▷ Multilevel mixed-effects models ▷ Nonlinear regression と操作
- Model タブ: Estimation method: Maximum likelihood (ML) （デフォルト）
  Dependent variable: `weight`
  Nonlinear expression: `{b1}+{U0[id]}+({b2}-{b1}-{U0[id]})*exp(-{b3}*time)`

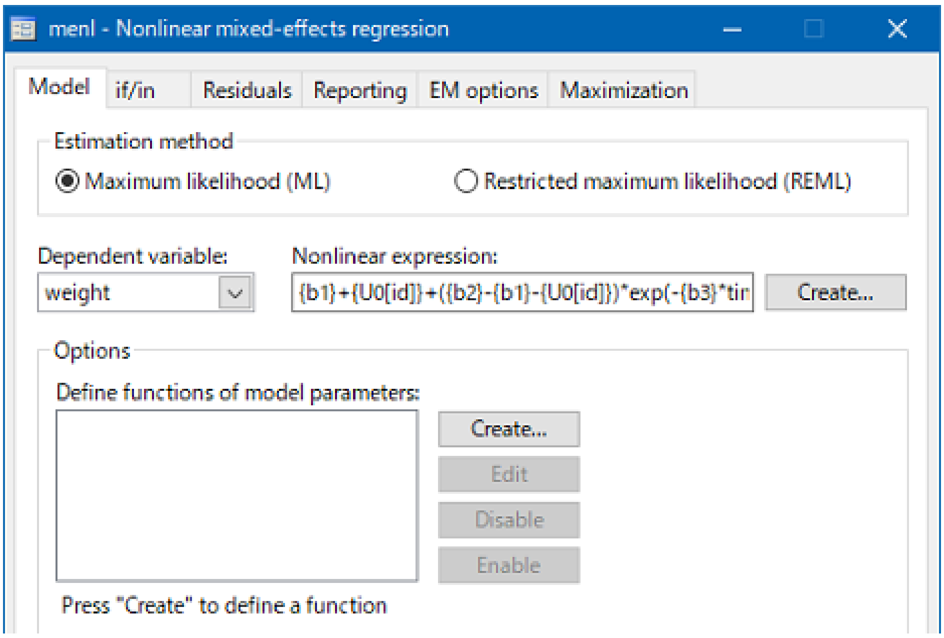

図 1 menl ダイアログ - Model タブ

```
. menl weight = {b1}+{U0[id]}+({b2}-{b1}-{U0[id]})*exp(-{b3}*time)

Obtaining starting values by EM:

Alternating PNLS/LME algorithm:

Iteration 1:    linearization log likelihood = -56.9757597

Computing standard errors:

Mixed-effects ML nonlinear regression           Number of obs     =        780
Group variable: id                              Number of groups  =         60

                                                Obs per group:
                                                              min =         13
                                                              avg =       13.0
                                                              max =         13

Linearization log likelihood =  -56.97576

------------------------------------------------------------------------------
      weight |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         /b1 |   4.707954   .1414511    33.28   0.000     4.430715    4.985193
         /b2 |   8.089432   .0260845   310.12   0.000     8.038307    8.140556
         /b3 |    4.13201   .0697547    59.24   0.000     3.995293    4.268726
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Identity                 |
                     var(U0) |   1.189809   .2180067      .8308316     1.70389
-----------------------------+------------------------------------------------
               var(Residual) |   .0439199   .0023148      .0396095    .0486995
------------------------------------------------------------------------------
```

特記事項：

1. {U0[id]} という表記によりグループ変数 id によって識別されるレベル（一角獣のレベル（レベル 2)）におけるランダム切片が規定されます。
2. ログは 3 つの部分から構成されます。
   (a) 初期値を改良するための EM 反復のログ。デフォルトではこのログは表示されませんが、emlog オプションを指定すればそれを表示させることができます。NLME モデルには複数の解が存在し、最大値 (global maximum) ではなく極大値 (local maximum) に収束してしまうことがあります。そのためいくつかの初期値を試してみることが大切です（セクション 3 参照)。
   (b) Lindstrom-Bates アルゴリズムからの線形化対数尤度値を示す反復ログ。“線形化”という用語は、対数尤度値が非線形の平均値関数を固定効果、変量効果に関し線形化することによって得られる線形の混合効果モデルに対応したものであることを意味します。

(c) “Computing standard errors”というメッセージ。これは最大化のための反復計算が終了し、標準誤差の算出が行われていることを示すものです。固定効果の推定値のみが欲しい場合には `nostderr` オプションを指定することによってこのステップをバイパスさせることも可能です。

3. “Mixed-effects ML nonlinear regression”というタイトルはモデルが最尤法によってフィットされたことを示しています。REML 法による推定値は `reml` オプションを指定することによって求めることができます。
4. ヘッダ情報は `mixed` からのものとほとんど変りませんが、すべての固定効果パラメータ（ただし定数項は除く）に対する結合有意性 (joint significance) の検定は非線形モデルの場合、意味を持たない可能性があるため、モデル全体に対する χ^2 統計量がレポートされてこないという点が `mixed` と異なります。
5. 最初のテーブルには固定効果の推定値が $\beta_1 = 4.71, \beta_2 = 8.09, \beta_3 = 4.13$ としてレポートされています。固定効果パラメータについてはその値が 0 と言えるか否かに関する z 検定の結果が自動的に出力されてきますが、非線形モデルの場合、このような 0 に対する検定は関心の対象とならないこともあるし、さらに言うなら適切とは言えないこともあります。意味のある検定は `test` コマンド（**[R] test** (*mwp-180*) 参照）によって実行することができます。
6. 2 番目のテーブルには分散成分 (variance components) の推定値が出力されます。最初のセクションには `id: Identity` というラベルが付加されています。これはモデルが `id` レベルの変量効果を有し、その分散共分散行列 $\mathbf{\Sigma}$ が単位行列 (identity matrix) — すべての変量効果が同一の分散を持つ — であることを示すものです。ここでのモデルの場合、モデル中には 1 つの変量効果しか含まれていないため、その共分散構造は意味を持ちません。出力中 `var(U0)` という表記でランダム切片の分散 σ_u^2 に関する情報が表示されています。それによると分散推定値は 1.19、その標準誤差は 0.22 であることがわかります。
7. `var(Residual)` と表記された行には全般的な誤差分散推定値が $\widehat{\mathrm{Var}}(\epsilon_{ij}) = \widehat{\sigma}_\epsilon^2 = 0.044$ と表示されています。

◁

▷ Example 2: **2 段階モデル**

評価版では割愛しています。

▷ Example 3: **共変量を含むモデル**

評価版では割愛しています。

▷ Example 4: **線形結合**

評価版では割愛しています。

▷ Example 5: **ランダム係数**

評価版では割愛しています。

5. 分散成分の検定

評価版では割愛しています。

▷ Example 6: **基本モデル**

評価版では割愛しています。

▷ Example 7: **分散成分の尤度比検定**

評価版では割愛しています。

▷ Example 8: Subject **内共変量**

評価版では割愛しています。

6. 変量効果の共分散構造

評価版では割愛しています。

▷ Example 9: **相関を伴う** 2 **階層モデル**

評価版では割愛しています。

▷ Example 10: **分散不均一性のチェック**

評価版では割愛しています。

7. 群内不均一分散誤差

評価版では割愛しています。

▷ Example 11: **べき乗型分散不均一性**

評価版では割愛しています。

▷ Example 12: **交互作用を含む分散不均一モデル**

評価版では割愛しています。

▷ Example 13: **予測**

評価版では割愛しています。

▷ Example 14: **群内誤差の相関構造**

評価版では割愛しています。

8. REML 推定法

評価版では割愛しています。

▷ Example 15: **薬物動態学モデル**

評価版では割愛しています。

▷ Example 16: **推定結果の解釈**

評価版では割愛しています。

9. 3 階層モデル

評価版では割愛しています。

▷ Example 17: 3 **階層モデル**

評価版では割愛しています。

▷ Example 18: AR(1) **誤差構造**

評価版では割愛しています。

▷ Example 19: **ブロック対角な共分散行列**

評価版では割愛しています。

10. 初期値の設定

評価版では割愛しています。

10.1 線形化アプローチ

評価版では割愛しています。

10.2 グラフィカルアプローチ

評価版では割愛しています。

10.3 回帰的アプローチ

評価版では割愛しています。

10.4 初期値の指定方法

評価版では割愛しています。

■