Reference manual of the nls2 library

Bouvier A., Huet S.
INRA, Unité de Biométrie
78352 Jouy-en-Josas Cedex
France

December 1993, last reviewed July 1998

Contents

Chapter 1

What is nls2

nls2 is a set of S functions and programs to estimate the parameters of a non-linear regression model over a given set of observations.

The regression function can be defined explicitly as a function of independent variables and of unknown parameters or it can be defined as the solution of a system of differential equations. Heteroscedasticity of errors can be taken into account by modelling the variance function.

Several additional tools are included: plotting functions, functions to analyze series of estimations, calculate confidence intervals and confidence regions for parameters and for functions of parameters, and functions to study calibration. The description of the models can be provided by using a symbolic syntax.

Chapter 2

Statistical and numerical methods

2.1 Parametric nonlinear regression model

The observed variable Y depends on m *independent variables* x through a known function f which depends on p unknown parameters θ. For each value of x, x_l, $(l = 1, \ldots, n)$, the observed *response* is denoted by Y_l. The errors $r_l = Y_l - f(x_l, \theta)$ are assumed to be independent. The variance is assumed constant, or dependent on x_l, θ, and/or on q unknown parameters β, through a known function v.

The following model is considered:

$$Y_l = f(x_l, \theta) + V_l^{1/2} \varepsilon_l, \quad \text{with } V_l = \sigma^2 v(x_l, \theta, \beta)/w_l,$$

where $(w_1, \ldots w_n)$ are known positive *weights*.

f is called the *regression function* and v the *variance function*.

Let $(\widehat{\theta}, \widehat{\beta})$ be the estimators of (θ, β). Some regularity conditions are required for their statistical properties (consistency, convergence in law) to be valid. For example (θ, β) lies in a compact subset of $R^p \times R^q$ (or more simply, each parameter varies in a bounded interval), f and v are twice continuously differentiable with respect to (θ, β).

nls2 offers several methods for estimating the parameters, depending on the assumptions on the error distribution. The parameters (θ, β) can be estimated simultaneously or alternatively. In addition, several other quantities are calculated: asymptotic variance matrix, residuals, ...

This manual gives only a brief description of the statistical and numerical methods used by **nls2**. Further details can be found in [3, 6, 9, 13, 1, 2, 18, 17, 5, 14, 10, 12, 4, 15], for example. Examples of pratical use are shown in [16].

A symbolic computations program is provided to simplify the description of functions f and v. From a symbolic description of f and v, it generates syntaxical trees and C-programs. The user can then choose between these two ways of evaluation.

It is the same thing when the function f is defined as the solution of a system of ordinary differential equations (or as a function of the solutions of the system): the user needs only describe the system of differential equations in a symbolic way. To carry out its numerical integration, **nls2** calls the program **lsoda** from the **ODEPACK** [11, 8] library.

2.2 Assumptions on the variance of the observations

Several estimators (or estimating methods) of (θ, β) are available. Choosing between them depends essentially on assumptions made on their variance and error distribution.

Note Parameter σ^2 is optional when describing the error variance. If $V_l = \sigma^2 v(x_l, \theta, \beta)/w_l$, then σ^2 is estimated by the residual variance or is assumed known. But σ^2 can also be estimated as a component of the vector of parameters β (see paragraph 2.7.1), in which case its asymptotic variance is computed, similarly to the other parameters.

Alternative assumptions on the variance function These assumptions will subsequently be refered to by using the key-words given in parentheses.

The variance of Y_l may be any of the following types:

- $VarY_l = \sigma^2$ (CST)
- $VarY_l = \sigma^2/w_l$ (SW)
- $VarY_l = \sigma^2 v(x_l, \theta)/w_l$ (VST)
- $VarY_l = v(x_l, \beta)/w_l$ (VB)
- $VarY_l = \sigma^2 v(x_l, \beta)/w_l$ (VSB)
- $VarY_l = \sigma^2 v(x_l, \theta, \beta)/w_l$ (VSTB)
- $VarY_l = v(x_l, \theta, \beta)/w_l$ (VTB)
- Case of a replicate experimental design (VI)

  For each distinct value of independent variables x and weights w, several *replications* (more than 2) of Y are observed. The model is then described in the following way:

  $$Y_{ij} = f(x_i, \theta) + V_i^{1/2}\varepsilon_{ij}$$

  i varying from 1 to k, j from 1 to n_i, where n_i is the number of replications of Y when $x = x_i$. The total number of observations is $n = \sum_{i=1,k} n_i$.

  $VarY_i$ may then be assumed to be equal to σ_i^2, which is unknown and estimated by the empirical variance (see paragraph 2.3.5).

2.3 Methods for estimating the parameters

Let $f_l = f(x_l, \theta)$, $\dfrac{\partial f_l}{\partial \theta}$, $\dfrac{\partial V_l}{\partial \theta}$, $\dfrac{\partial V_l}{\partial \beta}$, be the $p \times 1$ or $q \times 1$ vectors of derivatives of f_l and V_l with respect to θ and β, $l = 1, \ldots n$.

The estimating methods available in **nls2** are described in the following paragraphs. These methods will subsequently be refered to by using the key-words given in parentheses.

2.3.1 Estimation of $(\theta,\ \beta)$

- Maximum likelihood (MLTB)

  The error distribution is assumed to be gaussian.

$$(\widehat{\theta}, \widehat{\beta}) = Arg \min_{(\theta,\beta)} \frac{-2}{n} \log \mathcal{V}_n(\theta, \beta),$$

  where

$$\frac{-2}{n} \log \mathcal{V}_n(\theta, \beta) = \log(2\pi) + \frac{1}{n} \sum_{l=1}^{n} \log V_l + \frac{1}{n} \sum_{l=1}^{n} \frac{(Y_l - f_l)^2}{V_l}.$$

- Modified least squares (MLSTB)

  $(\widehat{\theta}, \widehat{\beta})$ is defined as the solution of the following system:

$$\begin{aligned} \sum_{l=1}^{n} \frac{1}{V_l} \frac{\partial f_l}{\partial \theta} (Y_l - f_l) &= 0 \\ \sum_{l=1}^{n} \frac{1}{V_l^2} \frac{\partial V_l}{\partial \beta} ((Y_l - f_l)^2 - V_l) &= 0 \end{aligned}$$

2.3.2 Estimation of θ

If β appears in the variance function, then β is set to a fixed value.

- Maximum likelihood (MLT)

  The error distribution is assumed to be gaussian.

$$\widehat{\theta} = Arg \min_{\theta} \frac{-2}{n} \log \mathcal{V}_n(\theta, \beta)$$

  where

$$\frac{-2}{n} \log \mathcal{V}_n(\theta, \beta) = \log(2\pi) + \frac{1}{n} \sum_{l=1}^{n} \log V_l + \frac{1}{n} \sum_{l=1}^{n} \frac{(Y_l - f_l)^2}{V_l}$$

- Weighted least squares (WLST)

$$\widehat{\theta} = Arg \min_{\theta} \mathcal{S}_n(\theta) = \frac{1}{n} \sum_{l=1}^{n} \frac{w_l}{v_l} (Y_l - f_l)^2$$

- Ordinary least squares (OLST)

$$\widehat{\theta} = Arg \min_{\theta} \mathcal{T}_n(\theta) = \frac{1}{n} \sum_{l=1}^{n} (Y_l - f_l)^2$$

- Modified least squares (MLST)

  $\widehat{\theta}$ is defined as the solution of the following system:

$$\sum_{l=1}^{n} \frac{1}{V_l} \frac{\partial f_l}{\partial \theta} (Y_l - f_l) = 0$$

- Weighted least squares, using the empirical variances as weights (VITWLS)

$$\widehat{\theta} = Arg \min_{\theta} \mathcal{U}_n(\theta) = \frac{1}{n} \sum_{i=1}^{k} \frac{1}{s_i^2} \sum_{j=1}^{n_i} (Y_{ij} - f_i)^2,$$

  where s_i^2 is the empirical variance calculated with replications (see paragraph 2.2, variance type VI):

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - Y_{i\cdot})^2, \tag{2.1}$$

  and where $Y_{i\cdot} = \sum_j Y_{ij}/n_i$, $i = 1, \dots k$, $j = 1, \dots n_i$.

  No specific assumption is made on V_i which is estimated by s_i^2.

2.3.3 Estimation of β

θ is set to a fixed value.

- Ordinary least squares (OLSB)

$$\widehat{\beta} = Arg \min_{\beta} \mathcal{R}_n(\beta) = \frac{1}{n} \sum_{l=1}^{n} \left((Y_l - f_l)^2 - V_l \right)^2$$

- Modified least squares (MLSB)

  $\widehat{\beta}$ is defined as the solution of the following system:

$$\sum_{l=1}^{n} \frac{1}{V_l^2} \frac{\partial V_l}{\partial \beta} ((Y_l - f_l)^2 - V_l) = 0$$

2.3.4 Estimation of $(\theta,\ \beta)$ in alternate steps

The parameters (θ, β) may be estimated alternatively. For example, θ is estimated by OLST in the first step, β is estimated by MLSB in the second step, then θ is estimated by MLST in the third step. When this option is chosen, it is not possible to estimate (θ, β) simultaneously in any step.

2.3.5 Estimation of σ^2

- If σ^2 appears in the description of the variance of Y, it may be estimated by different ways:
  - σ^2 is assumed to be known (KNOWN)
  - in case of a replicate experimental design
    - σ^2 may be estimated by $s^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - Y_{i\cdot})^2$. (VARREP)
    - or σ_i^2 is estimated by the empirical variance s_i^2 (VARINTRA)
  - σ^2 is estimated by the residual variance (RESID)

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{l=1}^{n} w_l \frac{(Y_l - f(x_l, \widehat{\theta}))^2}{v(x_l, \widehat{\theta}, \widehat{\beta})},$$

- When σ^2 does not occur in $Var(Y_l)$, σ^2 is ignored. (IGNORED)

2.4 Description of the estimating equations

The numerical estimation is in every case equivalent to solving a system called the *estimating equations*: $\mathcal{E}_n(\theta, \beta) = 0$, where $\mathcal{E}_n = B^T(\theta, \beta)\,(Z(Y) - \eta(\theta, \beta))$, in which $Z(Y)$ is the $nh \times 1$ vector of sufficient statistics, $\eta(\theta, \beta)$ is the expectation of $Z(Y)$, and $B(\theta, \beta)$ is a $nh \times (p+q)$ matrix.

Let $\dot{F}_\theta$ be the $n \times p$ matrix of derivatives of f with respect to θ, calculated at $x_1, \ldots x_n$: $(\dot{F}_\theta)_{la} = \frac{\partial f(x_l, \theta)}{\partial \theta_a}$, for $l = 1, \ldots n$, $a = 1, \ldots p$. Similarly, let $\dot{V}_\theta$ and $\dot{V}_\beta$ be the $n \times p$ and $n \times q$ matrices of derivatives of V with respect to θ et β.

Let $\Delta(\tilde{V})$ be the diagonal matrix with components $V_1, \ldots V_n$, and $\Delta(\tilde{f})$ the diagonal matrix with components $f_1, \ldots f_n$. Let $\tilde{f}$ be the $n \times 1$ vector with components $f_1, \ldots f_n$, $\tilde{f}^2$ the $n \times 1$ vector with components $f_1^2, \ldots f_n^2$, $\tilde{V}$ the $n \times 1$ vector with components $V_1, \ldots V_n$, and $\tilde{Y}$, $\tilde{Y}^2$ the $n \times 1$ vectors with components $Y_1, \ldots Y_n$ and $Y_1^2, \ldots Y_n^2$, respectively.

The estimating equations depend on the method used.

MLTB:

$$B^T = \begin{pmatrix} \dot{F}_\theta^T \Delta(\tilde{V})^{-1} - \dot{V}_\theta^T \Delta(\tilde{f})\Delta(\tilde{V})^{-2} & \dot{V}_\theta^T \Delta(\tilde{V})^{-2}/2 \\ -\dot{V}_\beta^T \Delta(\tilde{f})\Delta(\tilde{V})^{-2} & \dot{V}_\beta^T \Delta(\tilde{V})^{-2}/2 \end{pmatrix}$$

$$Z = \begin{pmatrix} \tilde{Y} \\ \tilde{Y}^2 \end{pmatrix} \quad \eta = \begin{pmatrix} \tilde{f} \\ \tilde{V} + \tilde{f}^2 \end{pmatrix}.$$

MLSTB:

$$B^T = \begin{pmatrix} \dot{F}_\theta^T \Delta(\tilde{V})^{-1} & 0 \\ -\dot{V}_\beta^T \Delta(\tilde{f})\Delta(\tilde{V})^{-2} & \dot{V}_\beta^T \Delta(\tilde{V})^{-2}/2 \end{pmatrix} \quad Z = \begin{pmatrix} \tilde{Y} \\ \tilde{Y}^2 \end{pmatrix} \quad \eta = \begin{pmatrix} \tilde{f} \\ \tilde{V} + \tilde{f}^2 \end{pmatrix}.$$

MLT:

$$B^T = \left(\begin{array}{cc} \dot{F}_\theta^T \Delta(\tilde{V})^{-1} - \dot{V}_\theta^T \Delta(\tilde{f})\Delta(\tilde{V})^{-2} & \dot{V}_\theta^T \Delta(\tilde{V})^{-2}/2 \end{array} \right)$$

$$Z = \left(\begin{array}{c} \tilde{Y} \\ \tilde{Y}^2 \end{array} \right) \quad \eta = \left(\begin{array}{c} \tilde{f} \\ \tilde{V} + \tilde{f}^2 \end{array} \right).$$

WLST:

$$B^T = \dot{F}_\theta^T \Delta(\tilde{V})^{-1} \quad Z = \tilde{Y} \quad \eta = \tilde{f}.$$

OLST:

$$B^T = \dot{F}_\theta^T \quad Z = \tilde{Y} \quad \eta = \tilde{f}.$$

MLST:

$$B^T = \dot{F}_\theta^T \Delta(\tilde{V})^{-1} \quad Z = \tilde{Y} \quad \eta = \tilde{f}.$$

VITWLS:

$$B^T = \dot{F}_\theta^T \Delta(\tilde{s}^2)^{-1} \quad Z = \tilde{Y} \quad \eta = \tilde{f},$$

where $\Delta(\tilde{s}^2)$ is the diagonal matrix with components $s_1^2, \dots s_n^2$.

OLSB:

$$B^T = \left(\begin{array}{cc} -\dot{V}_\beta^T \Delta(\tilde{f}) & \dot{V}_\beta^T/2 \end{array} \right) \quad Z = \left(\begin{array}{c} \tilde{Y} \\ \tilde{Y}^2 \end{array} \right) \quad \eta = \left(\begin{array}{c} \tilde{f} \\ \tilde{V} + \tilde{f}^2 \end{array} \right).$$

MLSB:

$$B^T = \left(\begin{array}{cc} -\dot{V}_\beta^T \Delta(\tilde{f})\Delta(\tilde{V})^{-2} & \dot{V}_\beta^T \Delta(\tilde{V})^{-2}/2 \end{array} \right) \quad Z = \left(\begin{array}{c} \tilde{Y} \\ \tilde{Y}^2 \end{array} \right) \quad \eta = \left(\begin{array}{c} \tilde{f} \\ \tilde{V} + \tilde{f}^2 \end{array} \right).$$

2.5 Asymptotic variance matrix

At the user's request, **nls2** calculates the *asymptotic variance matrix.*

Let D be the $nh \times (p+q)$ matrix of derivatives of η with respect to (θ, β): $D = (\dot{\eta}_\theta, \dot{\eta}_\beta)$, where $\dot{\eta}_\theta$ and $\dot{\eta}_\beta$ are the $nh \times p$ and $nh \times q$ matrices of derivatives of η with respect to θ and β.

Let $W = B^T D/n$, and let Var_Z be the variance matrix of Z:

- if $Z = \left(\begin{array}{c} \tilde{Y} \\ \tilde{Y}^2 \end{array} \right)$,

$$Var_Z = \left(\begin{array}{cc} \Delta(\tilde{V}) & \Delta(\tilde{\mu}_3) + 2\Delta(\tilde{f})\Delta(\tilde{V}) \\ \Delta(\tilde{\mu}_3) + 2\Delta(\tilde{f})\Delta(\tilde{V}) & \Delta(\tilde{\mu}_4) + 4\Delta(\tilde{f})^2\Delta(\tilde{V}) - \Delta(\tilde{V})^2 + 4\Delta(\tilde{f})\Delta(\tilde{\mu}_3) \end{array} \right).$$

  $\mu_{3,l}$ and $\mu_{4,l}$ are the third and fourth order moments of Y_l. $\tilde{\mu}_3$ is the vector with components $\mu_{3,1}, \dots \mu_{3,n}$ and $\Delta(\tilde{\mu}_3)$ the associated diagonal matrix. Notations for $\tilde{\mu}_4$ and $\Delta(\tilde{\mu}_4)$ are similar.

- if $Z = \tilde{Y}$, $Var_Z = \Delta(\tilde{V})$.

Then, the *asymptotic variance* matrix is:

$$AsVar_{\widehat{\theta},\widehat{\beta}} = \frac{1}{n} W^{-1} \frac{B^T Var_Z B}{n} W^{-1}. \tag{2.2}$$

For some estimators (for example, the maximum likelihood estimators), matrices W, B and Var_Z verify the following equation: $W^{-1} = BVar_Z B^T/n$. By extension, such estimators are called *efficient*. In this case, calculation of the asymptotic variance matrix simplifies to: $AsVar_{\widehat{\theta},\widehat{\beta}} = W^{-1}/n$, and there is no need to calculate Var_Z.

The asymptotic variance matrix is estimated by calculating W, B and Var_Z at $(\widehat{\theta}, \widehat{\beta})$. The third and fourth order moments can be estimated by the following different ways (these methods will subsequently be refered to by using the key-words given in parentheses).

- If the error distribution is assumed to be gaussian, (MUGAUSS)
  $\widehat{\mu}_{3,l} = 0$ and $\widehat{\mu}_{4,l} = 3\widehat{V}_l^2$, for $l = 1, \dots n$ where $\widehat{V}_l$ is an estimation of V_l.
  1. If $V_l = \sigma^2 v(x_l, \theta, \beta)/w_l$, then $\widehat{V}_l = \widehat{\sigma}^2 v(x_l, \widehat{\theta}, \widehat{\beta})/w_l$, where $\widehat{\sigma}^2$ can be calculated by different ways, depending on the user's choice (see paragraph 2.3.5).
  2. If $V_l = v(x_l, \theta, \beta)/w_l$, then $\widehat{V}_l = v(x_l, \widehat{\theta}, \widehat{\beta})/w_l$.
- The moments may be estimated using the residuals (MURES)
  Let $\widehat{r}_l = Y_l - f(x_l, \widehat{\theta})$:
  $$\widehat{\mu}_{3,l} = \frac{1}{n} \sum_l (\widehat{r}_l - \widehat{r}.)^3 \text{ and } \widehat{\mu}_{4,l} = \frac{1}{n} \sum_l (\widehat{r}_l - \widehat{r}.)^4,$$
  where $\widehat{r}.$ is the mean of $\widehat{r}_l$.
- or, in case of an experimental design with replications, (MURESREP)
  $$\widehat{\mu}_{3,i} = \frac{1}{n_i} \sum_j (\widehat{r}_{ij} - \widehat{r}_{i\cdot})^3 \text{ and } \widehat{\mu}_{4,i} = \frac{1}{n_i} \sum_j (\widehat{r}_{ij} - \widehat{r}_{i\cdot})^4,$$
  where $\widehat{r}_{ij} = Y_{ij} - f(x_i, \widehat{\theta})$.
- Finally, the moments may be known. (KNOWN)

When (θ, β) are estimated alternatively

When θ and β are estimated alternatively, θ is estimated in each odd step starting with the first step, β is estimated in even steps.

Let e be the number of the step. If e is odd, θ is estimated by $\widehat{\theta}_e$ and if e is even, β is estimated by $\widehat{\beta}_e$. $\widehat{\theta}_e$ or $\widehat{\beta}_e$ is the solution of the estimating equations $B_e^T(Z_e(Y) - \eta_e) = 0$. The dimension of Z_e is $nh_e \times 1$, and the dimension of B_e^T is $p \times nh_e$ if e is odd, and $q \times nh_e$ if e is even.

Let N_e be the number of the steps requested.

In the first step β is assigned an arbitrary value β_0. θ is estimated by $\widehat{\theta}_1$ the solution of:

$$B_1^T(\theta, \beta_0)(Z_1(Y) - \eta_1(\theta, \beta_0)) = 0.$$

The asymptotic variance matrix of $\widehat{\theta}_1$ is defined as before by:

$$AsVar_{\widehat{\theta}_1} = \frac{1}{n} W_1^{-1}(\widehat{\theta}_1, \beta_0) \frac{B_1^T(\widehat{\theta}_1, \beta_0) Var_{Z_1}(\widehat{\theta}_1, \beta_0) B_1(\widehat{\theta}_1, \beta_0)}{n} W_1^{-1}(\widehat{\theta}_1, \beta_0).$$

Example let $\widehat{\theta}_1$ the OLST estimator. Let $\dot{F}_{\theta,1}$ be the value of matrix $\dot{F}_\theta$ at $\widehat{\theta}_1$, and $\tilde{V}_{1,0}$ the vector with components V_l calculated at $(\widehat{\theta}_1, \beta_0)$. Then:

$$AsVar(\widehat{\theta}_1) = \frac{1}{n} \left(\dot{F}_{\theta,1}^T \dot{F}_{\theta,1} \right)^{-1} \frac{\dot{F}_{\theta,1}^T \Delta(\tilde{V}_{1,0}) \dot{F}_{\theta,1}}{n} \left(\dot{F}_{\theta,1}^T \dot{F}_{\theta,1} \right)^{-1}.$$

In the second step β is estimated by $\widehat{\beta}_2$ defined as the solution of:

$$B_2^T(\widehat{\theta}_1, \beta)(Z_2(Y) - \eta_2(\widehat{\theta}_1, \beta)) = 0.$$

Let $\Sigma_{\widehat{\beta}_2}$ be the asymptotic variance matrix of $\widehat{\beta}_2$ which would be calculated if $\theta = \widehat{\theta}_1$ were known:

$$\Sigma_{\widehat{\beta}_2} = \frac{1}{n} W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2) \frac{B_2^T(\widehat{\theta}_1, \widehat{\beta}_2) Var_{Z_2}(\widehat{\theta}_1, \widehat{\beta}_2) B_2(\widehat{\theta}_1, \widehat{\beta}_2)}{n} W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2).$$

Let $\Sigma_{\widehat{\theta}_1}$ be the asymptotic variance matrix of $\widehat{\theta}_1$ calculated at $(\widehat{\theta}_1, \widehat{\beta}_2)$:

$$\Sigma_{\widehat{\theta}_1} = \frac{1}{n} W_1^{-1}(\widehat{\theta}_1, \widehat{\beta}_2) \frac{B_1^T(\widehat{\theta}_1, \widehat{\beta}_2) Var_{Z_1}(\widehat{\theta}_1, \widehat{\beta}_2) B_1(\widehat{\theta}_1, \widehat{\beta}_2)}{n} W_1^{-1}(\widehat{\theta}_1, \widehat{\beta}_2).$$

Let $W_{2,\theta}$ be the $q \times p$ matrix defined by $W_{2,\theta} = B_2^T \dot{\eta}_{2,\theta}/n$, and Σ_{12} the $q \times p$ matrix defined by $\Sigma_{12} = B_2^T Cov_{Z_2,Z_1} B_1/n$, where Cov_{Z_2,Z_1} is the $nh_2 \times nh_1$ variance matrix matrix between Z_2 and Z_1.

Then the asymptotic variance matrix of $\widehat{\beta}_2$, is:

$$AsVar_{\widehat{\beta}_2} = \Sigma_{\widehat{\beta}_2} + $$
$$W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2) \left(W_{2,\theta}(\widehat{\theta}_1, \widehat{\beta}_2) \Sigma_{\widehat{\theta}_1} - \frac{2}{n} \Sigma_{12}(\widehat{\theta}_1, \widehat{\beta}_2) W_1^{-1}(\widehat{\theta}_1, \widehat{\beta}_2) \right) W_{2,\theta}(\widehat{\theta}_1, \widehat{\beta}_2) W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2)$$

Example let $\widehat{\beta}_2$ be the MLSB estimator and assume that the third and fourth moments of Y are those of a gaussian variable. $\Sigma_{\widehat{\beta}_2} = \frac{1}{n} W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2)$, with:

$$\begin{aligned} W_2^{-1}(\widehat{\theta}_1, \widehat{\beta}_2) &= \frac{1}{2n} \dot{V}_{\beta,1,2}^T \Delta(\tilde{V}_{1,2})^{-2} \dot{V}_{\beta,1,2} \\ \Sigma_{12} &= 0 \\ W_{2,\theta}(\widehat{\theta}_1, \widehat{\beta}_2) &= \frac{1}{2n} \dot{V}_{\beta,1,2}^T \Delta(\tilde{V_{1,2}})^{-2} \dot{V}_{\theta,1,2}, \end{aligned}$$

where $\dot{V}_{\beta,1,2}^T$ and $\dot{V}_{\theta,1,2}$ are the matrices $\dot{V}_\beta^T$ et $\dot{V}_\theta^T$ calculated at $(\widehat{\theta}_1, \widehat{\beta}_2)$, $\tilde{V}_{1,2}$ is the vector with components V_l calculated at $(\widehat{\theta}_1, \widehat{\beta}_2)$.

In the third step θ is estimated by $\widehat{\theta}_3$ defined as the solution of:

$$B_3^T(\theta,\widehat{\beta}_2)(Z_3(Y)-\eta_3(\theta,\widehat{\beta}_2))=0.$$

Let $\Sigma_{\widehat{\theta}_3}$ be the asymptotic variance matrix of $\widehat{\theta}_3$ which would be calculated if $\widehat{\beta}_2$ were known:

$$\Sigma_{\widehat{\theta}_3}=\frac{1}{n}W_3^{-1}(\widehat{\theta}_3,\widehat{\beta}_2)\frac{B_3^T(\widehat{\theta}_3,\widehat{\beta}_2)Var_{Z_3}(\widehat{\theta}_3,\widehat{\beta}_2)B_3(\widehat{\theta}_3,\widehat{\beta}_2)}{n}W_3^{-1}(\widehat{\theta}_3,\widehat{\beta}_2).$$

If η_3 does not depend on β, which is the case for estimators WLST, OLST, MLST, then the asymptotic variance matrix of $\widehat{\theta}_3$ is $\Sigma_{\widehat{\theta}_3}$ and does not depend on the estimating methods chosen in steps 1 and 2.

If not, let $\Sigma_{\widehat{\beta}_2}$ be the asymptotic variance of $\widehat{\beta}_2$ which would be calculated if $\theta=\widehat{\theta}_3$ were known:

$$\Sigma_{\widehat{\beta}_2}=\frac{1}{n}W_2^{-1}(\widehat{\theta}_3,\widehat{\beta}_2)\frac{B_2^T(\widehat{\theta}_3,\widehat{\beta}_2)Var_{Z_2}(\widehat{\theta}_3,\widehat{\beta}_2)B_2(\widehat{\theta}_3,\widehat{\beta}_2)}{n}W_2^{-1}(\widehat{\theta}_3,\widehat{\beta}_2).$$

Let $W_{3,\beta}$ be the $p\times q$ matrix defined by $W_{3,\beta}=B_3^T\dot{\eta}_{3,\beta}/n$, and as before, $W_{2,\theta}=B_2^T\dot{\eta}_{2,\theta}/n$, $W_1(\widehat{\theta}_3,\widehat{\beta}_2)=B_1^TD_1/n$.

Let Σ_{13} be the $p\times p$ matrix defined by $\Sigma_{13}=B_3^TCov_{Z_3,Z_1}B_1/n$ and Σ_{23} the $p\times q$ matrix, $\Sigma_{23}=B_3^TCov_{Z_3,Z_2}B_2/n$.

Then, the asymptotic variance matrix of $\widehat{\theta}_3$ is written as:

$$AsVar_{\widehat{\theta}_3}=\Sigma_{\widehat{\theta}_3}+W_3^{-1}\left(W_{3,\beta}\Sigma_{\widehat{\beta}_2}+\frac{2}{n}(G_{13}-\Sigma_{23})W_2^{-1}\right)W_{3,\beta}W_3^{-1},$$

where $G_{13}=\Sigma_{13}W_1^{-1}W_{2,\theta}$ (these matrices are calculated at $(\widehat{\theta}_3,\widehat{\beta}_2)$).

Example let $\widehat{\theta}_3$ be the MLST estimator. η_3 does not depend on β. Thus,

$$AsVar_{\widehat{\theta}_3}=\frac{1}{n}\left(\frac{1}{n}\dot{F}_{\theta,3}^T\Delta(\tilde{V}_{3,2})^{-1}\dot{F}_{\theta,3}\right)^{-1}\frac{\dot{F}_{\theta,3}^T\dot{F}_{\theta,3}}{n}\left(\frac{1}{n}\dot{F}_{\theta,3}^T\Delta(\tilde{V}_{3,2})^{-1}\dot{F}_{\theta,3}\right)^{-1},$$

with the same notations as for the preceding steps.

2.6 Numerical method

The algorithm used to solve the estimating equations is based on the *Gauss-Newton* algorithm. The *Gauss-Marquardt* modification is also available.

Let (θ_{it},β_{it}) be the current value of (θ,β) at the beginning of iteration number *it*. A new value $(\theta_{it+1},\beta_{it+1})$ is calculated at each iteration of the algorithm.

2.6.1 The Gauss-Newton algorithm

This algorithm is based on a linear approximation of $\eta(\theta, \beta)$:

$$\eta(\theta, \beta) = \eta(\theta_{it}, \beta_{it}) + D(\theta_{it}, \beta_{it}) \begin{pmatrix} \theta - \theta_{it} \\ \beta - \beta_{it} \end{pmatrix},$$

where $(\theta_{it}, \beta_{it})$ is the current value of (θ, β).

Transposing this approximation into the estimating equations leaves us with the following linear system:

$$B^T(\theta_{it}, \beta_{it}) D(\theta_{it}, \beta_{it}) \begin{pmatrix} \theta - \theta_{it} \\ \beta - \beta_{it} \end{pmatrix} = B^T(\theta_{it}, \beta_{it})(Z(Y) - \eta(\theta_{it}, \beta_{it})).$$

which has to be solved for (θ, β).

The solution of this system is the Gauss-Newton approximation. Let δ_{it} be the *new direction*:

$$\delta_{it} = \frac{1}{n} W^{-1}(\theta_{it}, \beta_{it}) B^T(\theta_{it}, \beta_{it})(Z(Y) - \eta(\theta_{it}, \beta_{it})),$$

then $(\theta,\ \beta) = (\theta_{it},\ \beta_{it}) + \delta_{it}^T$ is the solution. One should check that these new parameter values minimize the *fitting criterion* if the estimator was defined by the minimization of a criterion. In all other cases the estimator was defined as the solution of the estimating equations, $\mathcal{E}_n(\theta, \beta) = 0$, and the criterion is:

$$\mathcal{Q}_n(\theta, \beta) = (Z(Y) - \eta)^T \frac{B^T}{n} W^{-1} \frac{B}{n} (Z(Y) - \eta). \tag{2.3}$$

The fitting criterion is denoted by $\mathcal{C}_n$ where $\mathcal{C}_n$ is one of the following criterions: $-2\log \mathcal{V}_n/n$, $\mathcal{S}_n$, $\mathcal{T}_n$, $\mathcal{R}_n$, $\mathcal{Q}_n$.

$\mathcal{C}_n$ is calculated for three values of the parameters: $(\theta_{it}, \beta_{it})$, $(\theta_{it+\delta},\ \beta_{it+\delta}) = (\theta_{it},\ \beta_{it}) + \delta_{it}^T$, and $(\theta_{it+\delta/2},\ \beta_{it+\delta/2}) = (\theta_{it},\ \beta_{it}) + \delta_{it}^T/2$.

The *optimal step* is calculated by the following way:

- if $\mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, then $(\theta_{it+1}, \beta_{it+1}) = (\theta_{it+\delta}, \beta_{it+\delta})$.
- if $\mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, then the optimal step ω_{it} is calculated using a quadratic approximation of these 3 points, and $(\theta_{it+1},\ \beta_{it+1}) = (\theta_{it},\ \beta_{it}) + \omega_{it}\delta_{it}^T$.
- if $\mathcal{C}_n(\theta_{it}, \beta_{it}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, then $(\theta_{it+\delta},\ \beta_{it+\delta})$ is calculated again: $(\theta_{it+\delta},\ \beta_{it+\delta}) = (\theta_{it},\ \beta_{it}) + \omega_c\delta_{it}^T$ where ω_c has a given value. This correction is not allowed more than N_c times.
- if $\mathcal{C}_n(\theta_{it}, \beta_{it}) = \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}) = \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})$, then if the estimating method is not modified least squares, $(\theta_{it+\delta},\ \beta_{it+\delta})$ is calculated again: $(\theta_{it+\delta},\ \beta_{it+\delta}) = (\theta_{it},\ \beta_{it}) + \omega_c\delta_{it}^T$ where ω_c has a given value. This correction is not allowed more than N_c times.

  If the estimating method is modified least squares, see the paragraph 2.6.3.

2.6.2 Gauss-Marquardt algorithm

The Gauss-Marquardt algorithm is a modification of the Gauss-Newton algorithm: a matrix λI is added to W at each iteration, where I is the identity matrix with same dimensions as W, and λ is a positive scalar. This modification improves the numerical stability, especially when the starting values of the parameters are far from the solution.

The direction is then equal to:

$$\delta_{it} = \frac{1}{n} \left(W(\theta_{it}, \beta_{it}) + \lambda I\right)^{-1} B^T(\theta_{it}, \beta_{it}) \left(Z(Y) - \eta(\theta_{it}, \beta_{it})\right).$$

$\mathcal{C}_n$ is calculated for three values of the parameters: $(\theta_{it}, \beta_{it})$, $(\theta_{it+\delta},\ \beta_{it+\delta}) = (\theta_{it},\ \beta_{it}) + \delta_{it}^T$, and $(\theta_{it+\delta/2},\ \beta_{it+\delta/2}) = (\theta_{it},\ \beta_{it}) + \delta_{it}^T/2$.

The *optimal step* is calculated by the following way:

- if $\mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, then $(\theta_{it+1}, \beta_{it+1}) = (\theta_{it+\delta}, \beta_{it+\delta})$.
- if $\mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, then the optimal step ω_{it} is calculated using a quadratic approximation of these 3 points, and $(\theta_{it+1},\ \beta_{it+1}) = (\theta_{it},\ \beta_{it}) + \omega_{it}\delta_{it}^T$.

  In these two cases, the value of λ decreased: λ is multiplied by a given value λ_1 $(0 < \lambda_1 < 1)$.
- if $\mathcal{C}_n(\theta_{it}, \beta_{it}) = \inf\{\mathcal{C}_n(\theta_{it}, \beta_{it}), \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}), \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})\}$, the direction is recalculated with an increased value of λ: λ is multiplied by a given value λ_2 $(\lambda_2 > 1)$. This correction is not carried out more than N_c times.
- if $\mathcal{C}_n(\theta_{it}, \beta_{it}) = \mathcal{C}_n(\theta_{it+\delta/2}, \beta_{it+\delta/2}) = \mathcal{C}_n(\theta_{it+\delta}, \beta_{it+\delta})$, then if the estimating method is not modified least squares, and if λ is less than a given maximal value, the direction is recalculated with an increased value of λ: λ is multiplied by a given value λ_2 $(\lambda_2 > 1)$. This correction is not carried out more than N_c times.

  Otherwise, see the paragraph 2.6.3.

2.6.3 Stopping the process

The iterative process begins with *starting values of the parameters*, (θ_0, β_0) and is pursued until convergence, or occurrence of an error.

Convergence of the algorithm The process is stopped when a *stopping criterion* is lower than a value fixed "a priori", denoted by $\mathcal{Q}_{stop}$. This stopping criterion is the norm of the estimating equations, denoted by $\mathcal{Q}_n$, see equation (2.3).

In case of Gauss-Marquardt algorithm, the value of λ must be lower than an upper bound denoted by λ_{max}.

In case of a modified least squares estimator criterion, $\mathcal{Q}_n = \mathcal{C}_n$. If the three values of $\mathcal{C}_n$ obtained when calculating the *optimal step* are equal, and if the two conditions above are fulfilled, then the process is stopped.

Stopping the process with errors

- The model calculation, i.e the calculation of functions f and v, may be impossible starting from $(\theta, \beta) = (\theta_0, \beta_0)$. The user has to choose new starting values.

- The model may be impossible to calculate for the value $(\theta_{it+\delta}, \beta_{it+\delta})$ of the parameters, in which case the direction is modified so that new values of the model can be calculated:

$$(\theta_{it+\delta},\ \beta_{it+\delta}) = (\theta_{it},\ \beta_{it}) + \omega_{err}\delta_{it},$$

  where ω_{err} is a positive given value. At most N_{err} trials are allowed.

- A direction that allows the fitting criterion minimization cannot be found after N_c trials.

- The number of iterations has reached a value fixed "a priori" denoted by I_{max}.

2.6.4 Note on matrix inversion

For the modified least squares estimator of (θ, β), matrix W is block symmetric, i.e:

$$W = \begin{pmatrix} P & 0 \\ Q & R \end{pmatrix}$$

where P is a $p \times p$ symmetric matrix, R is a $q \times q$ symmetric matrix, Q is a $q \times p$ matrix and 0 a $p \times q$ matrix with components 0. W^{-1} is then calculated by the following way:

$$W^{-1} = \begin{pmatrix} P^{-1} & 0 \\ -R^{-1}QP^{-1} & R^{-1} \end{pmatrix}.$$

Matrix W is symmetric for all the other available estimators.

2.7 Taking into account constraints on the parameters

The estimation under numerical constraints on the parameters or under some equality constraints between parameters, and the estimation for several curves (see paragraph 2.7.2), are possible without any modification on the model. Until now, θ and β were the set of parameters that need to be estimated. These are called *active parameters.* In what follows, we shall introduce some new terms to make the different types of parameters clearer.

2.7.1 Basic parameters

Basic parameters are the parameters that describe the model. They occur in the definition of the functions f and v. Let p_{basic} and q_{basic} be the number of basic parameters in f and v, respectively.

Examples

1. f is a logistic function, and $\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3, \bar{\theta}_4$ are the basic parameters:

$$f(x, \bar{\theta}) = \bar{\theta}_1 + \frac{\bar{\theta}_2}{1 + \exp(\bar{\theta}_3 + \bar{\theta}_4 x)}.$$

The model is $Y = f(x, \bar{\theta}) + \varepsilon$, with $Var(\varepsilon) = \sigma^2$ $(v(x) = 1)$. Then $p_{basic} = 4$ and $q_{basic} = 0$.

2. If $Var(\varepsilon) = \sigma^2 f(x, \bar{\theta})^{\bar{\beta}}$ $(v(x, \bar{\theta}, \bar{\beta}) = f(x, \bar{\theta})^{\bar{\beta}})$, where $\bar{\beta}$ is an unknown parameter to be estimated, then $q_{basic} = 1$.

In these two examples, σ^2 might be known, or might be estimated by the residual variance, or by the variance calculated with replications (see paragraph 2.3.5).

3. It is also possible to introduce σ^2 into the parameters defining the variance function: $Var(\varepsilon) = v(x, \bar{\theta}, \bar{\beta}) = \bar{\beta}_1 f(x, \bar{\theta})^{\bar{\beta}_2}$. In that case, $(\bar{\beta}_1, \bar{\beta}_2)$ are estimated simultaneously and $q_{basic} = 2$.

2.7.2 Multiple parameters

In some cases, people are interested in estimating the parameters for different sets of data, corresponding to different sets of parameters in the same model. A *curve* is associated to each set of data.

In case of several curves, let us say c curves, the *multiple parameters* are the basic parameters repeated c times. The number of multiple parameters is $p_{mult} = c \times p_{basic}$, and $q_{mult} = c \times q_{basic}$.

When σ^2 appears in the definition of the variance function, independently of the description of v (see example **2.**), σ^2 is assumed to be identical for all curves.

2.7.3 Distinct parameters

Equality constraints between parameters can be introduced. The numbers of *distinct parameters*, p_{dist} and q_{dist}, are the number of *multiple parameters* minus the number of equality constraints.

Examples

4. Let us consider example **1.** with 2 curves. If the curves are assumed distinct, then $p_{dist} = p_{mult} = 2 \times p_{basic} = 8$ and $q_{dist} = 0$. The vector of distinct parameters is then

$$\left(\bar{\theta}_1^1, \bar{\theta}_2^1, \bar{\theta}_3^1, \bar{\theta}_4^1, \bar{\theta}_1^2, \bar{\theta}_2^2, \bar{\theta}_3^2, \bar{\theta}_4^2\right)$$

where the indice identifies the parameter, and the exponent identifies the curve.

5. If the asymptotes $\bar{\theta}_1$ and $\bar{\theta}_2$ are assumed equal, $p_{dist} = 6$, and the vector of distinct parameters is

$$\left(\bar{\theta}_1^1, \bar{\theta}_2^1, \bar{\theta}_3^1, \bar{\theta}_4^1, \bar{\theta}_3^2, \bar{\theta}_4^2\right),$$

the vector of multiple parameters is

$$\left(\bar{\theta}_1^1, \bar{\theta}_2^1, \bar{\theta}_3^1, \bar{\theta}_4^1, \bar{\theta}_1^1, \bar{\theta}_2^1, \bar{\theta}_3^2, \bar{\theta}_4^2\right).$$

2.7.4 Active parameters

Different types of *numerical constraints* can be assigned to the parameters:

- The value of one or several parameters is fixed.
- One or several parameters must satisfy inequality constraints. In that case, a transformation of their current value is carried out. Let P be one of the *distinct parameters* of the regression or variance function.

  If $P > b_{\min}$, the transformation $\zeta = \sqrt{P - b_{\min}}$ is used.

  If $P < b_{\max}$, the transformation $\zeta = \sqrt{b_{\max} - P}$ is used

  If $b_{\min} < P < b_{\max}$ the transformation $\zeta = \arcsin\sqrt{\dfrac{P - b_{\min}}{b_{\max} - b_{\min}}}$ is used.

The total number of parameters to estimate, denoted by *active parameters* , equals the number of *distinct parameters* minus the number of numerical equality constraints. The *active parameters* are the *distinct parameters* transformed in order to take into account the inequality constraints.

The numerical processing is carried out in the *active parameters* space.

Examples

6. Let us again consider example **3.** and assume that the asymptote $\bar{\theta}_1$ is zero, and that the parameter $\bar{\beta}_1$ is positive. Then $p_{active} = 3$, $q_{active} = 2$; the active parameters appearing in the regression function are $(\bar{\theta}_2, \bar{\theta}_3, \bar{\theta}_4)$, and in the variance function $(\zeta_{\beta,1}, \bar{\beta}_2)$, where $\zeta_{\beta,1} = \sqrt{\bar{\beta}_1}$.

2.8 Definition of the model

The values of f, v and of their derivatives are calculated at each iteration. Symbolic computations software is provided in order to save the user writing a routine which calculates these values.

f may either be explicitly defined (see example **1.**), or may be a function of the solution of an ordinary differential equations system, denoted by *odes*.

Let t be the integration variable (time for example) and consider the following *odes* with N_{eq} equations:

$$\begin{aligned}\frac{dF_\nu}{dt} &= A_\nu(t, F_1, \ldots F_{N_{eq}})\\ F_\nu(t_0) &= F_{\nu,0}\end{aligned}$$

for $\nu = 1, \ldots N_{eq}$. Each of the functions F_ν may depend on the independent variables x (this implies that the system to be integrated is different for each value of x), and on unknown parameters, denoted by θ_{odes}. Let $F_{\nu,0}$ be the *initial values of the system.* $F_{\nu,0}$ may be known or unknown, in which case they are considered as parameters in the model and are denoted by θ_{ci}.

Let T be a $n \times J$ matrix of t values, and let $F_\nu(T_{l,\ell}, x_l, \theta_{odes}, \theta_{ci})$, for $\nu = 1, \ldots N_{eq}$, be the solution of the system calculated at $T_{l,\ell}$, and possibly x_l (if the system depends on x) for $\ell = 1, \ldots J$ and $l = 1, \ldots n$. Thus:

$$f(x_l, \theta) = \phi\left(F_\nu(T_{l,\ell}, x_l, \theta_{odes}, \theta_{ci}), \nu = 1, \ldots N_{eq}, \ell = 1, \ldots J; x_l, \theta_\phi\right),$$

where θ_ϕ is the vector of parameters θ appearing in ϕ.

Examples

7. The integration variable is the independent variable; the model is a model with 2 compartments where the observed variable is the amount of product in the first compartment. The variations of this variable are modelled as a function of time. The initial values of the system are known.

$$\begin{aligned}\frac{dF_1}{dt} &= -\bar\theta_1 F_1 + \bar\theta_2 F_2\\ \frac{dF_2}{dt} &= -\bar\theta_3 F_1 + (\bar\theta_2 + \bar\theta_4) F_2\\ F_1(t_0) &= F_{1,0}\\ F_2(t_0) &= F_{2,0}.\end{aligned}$$

$N_{eq} = 2$, $T_{l,1} = x_l$, et $J = 1$. $\phi(F_1(x_l), F_2(x_l)) = F_1(x_l)$. The set of parameters to be estimated $\theta = \theta_{odes}$ is $(\bar\theta_1, \bar\theta_2, \bar\theta_3, \bar\theta_4)$.

8. The integration variable is not an independent variable; the model is a model with 2 compartments where the observed variable depends on the amount of product in the first compartment at time T_{chosen} and on another variable x (temperature, for example). The initial conditions of the system are parameters to be estimated.

$$\begin{aligned}\frac{dF_1}{dt} &= -\bar\theta_3 F_1 + \bar\theta_4 F_2\\ \frac{dF_2}{dt} &= -\bar\theta_5 F_1 + (\bar\theta_4 + \bar\theta_6) F_2\\ F_1(t_0) &= \bar\theta_1\\ F_2(t_0) &= \bar\theta_2.\end{aligned}$$

$N_{eq} = 2$, $T(l,1) = T_{chosen}$, for all l. $J = 1$, and

$$\phi(F_1(T_{chosen}), F_2(T_{chosen}), x_l) = F_1(T_{chosen})(1 + \exp(\bar{\theta}_7 x_l)).$$

The set of parameters to be estimated $\theta = (\theta_{ci}, \theta_{odes}, \theta_\phi)$, where $\theta_{ci} = (\bar{\theta}_1, \bar{\theta}_2)$, $\theta_{odes} = (\bar{\theta}_3, \bar{\theta}_4, \bar{\theta}_5, \bar{\theta}_6)$, and $\theta_\phi = \bar{\theta}_7$.

Calculation of derivatives with respect to θ The derivatives of f with respect to the parameters are calculated as in the case where f is explicitly defined:

$$\frac{\partial f}{\partial \theta_a} = \frac{\partial \phi}{\partial \theta_a} + \sum_{\nu=1}^{N_{eq}} \frac{\partial \phi}{\partial F_\nu} \frac{\partial F_\nu}{\partial \theta_a} \quad \text{for } a = 1, \dots p.$$

nls2 uses program **lsoda** from the **ODEPAK** library for calculating the values of F_ν and $\frac{\partial F_\nu}{\partial \theta}$. The *odes* is the set of equations defining the F_ν and their derivatives with respect to the parameters.

Description of the system to be integrated

When the initial values of the system are known $F = (F_1, \dots F_{N_{eq}})$ and $\partial F / \partial \theta^a_{odes} = (\partial F_1 / \partial \theta^a_{odes}, \dots \partial F_{N_{eq}} / \partial \theta^a_{odes})$ must be calculated.
Let p_{odes} be the number of θ_{odes} parameters. The following system with $N_{eq}(1 + p_{odes})$ equations must be solved for each value of x and θ_{odes}:

$$\begin{aligned}
\frac{dF_\nu}{dt} &= A_\nu(t, F_1, \dots F_{N_{eq}}) \\
\frac{d}{dt}\frac{\partial F_\nu}{\partial \theta^a_{odes}} &= \frac{\partial A_\nu}{\partial \theta^a_{odes}}(t, F_1, \dots F_{N_{eq}}) + \sum_{\nu'=1}^{N_{eq}} \frac{\partial A_\nu}{\partial F_{\nu'}}(t, F_1, \dots F_{N_{eq}}) \frac{\partial F_{\nu'}}{\partial \theta^a_{odes}}.
\end{aligned}$$

The initial values of the system are:

$$\begin{aligned}
F_\nu(t_0) &= F_{\nu,0} \\
\frac{\partial F_\nu}{\partial \theta^a_{odes}}(t_0) &= 0
\end{aligned}$$

where $a = 1, \dots p_{odes}$ and t_0 is the initial value of t.

When the initial values of the system are parameters to be estimated In this case, F_ν, $\partial F_\nu / \partial \theta^a_{odes}$ and $\partial F_\nu / \partial \theta^b_{ci}$ must be calculated by integrating the following system with $N_{eq}(1 + p_{odes} + N_{eq})$ equations:

$$\begin{aligned}
\frac{dF_\nu}{dt} &= A_\nu(t, F_1, \dots F_{N_{eq}}) \\
\frac{d}{dt}\frac{\partial F_\nu}{\partial \theta^a_{odes}} &= \frac{\partial A_\nu}{\partial \theta^a_{odes}}(t, F_1, \dots F_{N_{eq}}) + \sum_{\nu'=1}^{N_{eq}} \frac{\partial A_\nu}{\partial F_{\nu'}}(t, F_1, \dots F_{N_{eq}}) \frac{\partial F_{\nu'}}{\partial \theta^a_{odes}} \\
\frac{d}{dt}\frac{\partial F_\nu}{\partial \theta^b_{ci}} &= \sum_{\nu'=1}^{N_{eq}} \frac{\partial A_\nu}{\partial F_{\nu'}}(t, F_1, \dots F_{N_{eq}}) \frac{\partial F_{\nu'}}{\partial \theta^b_{ci}}.
\end{aligned}$$

The initial values are:

$$
\begin{aligned}
F_\nu(t_0) &= \theta_{ci,\nu} \\
\frac{\partial F_\nu}{\partial \theta^a_{odes}}(t_0) &= 0 \\
\frac{\partial F_\nu}{\partial \theta_{ci,\nu}}(t_0) &= 1 \\
\frac{\partial F_\nu}{\partial \theta_{ci,\nu'}}(t_0) &= 0 \quad si \;\; \nu' \neq \nu
\end{aligned}
$$

where t_0 is the initial value of t, and where ν, ν' vary from 1 to N_{eq} and a, b from 1 to p_{odes}.

Chapter 3

List of notations

3.1 Key words

active parameters: see paragraph 2.7.4.

asymptotic variance: see paragraph 2.5.

constraints: see paragraph 2.7.

curves: see paragraphs 2.7 and 2.7.2.

basic parameters: see paragraph 2.7.1.

direction denoted by δ_{it} at iteration it. See paragraph 2.6.

distinct parameters: see paragraph 2.7.3.

efficient: an estimator is *efficient* when the matrices W, B and Var_Z verify: $W^{-1} = BVar_Z B^T/n$. See page 9.

equality constraints between parameters: see paragraph 2.7.3.

estimating equations: $\mathcal{E}_n(\theta, \beta)$ See page 7.

fitting criterion: $\mathcal{C}_n(\theta, \beta)$. See page 12.

Gauss-Marquardt: algorithm for estimating the parameters. See paragraph 2.6.2.

Gauss-Newton: algorithm for estimating the parameters. See paragraph 2.6.1.

independent variable: X. See paragraph 2.1.

initial values of the system: see paragraph 2.8.

integration values: T. See paragraph 2.8.

integration variable: t. See paragraph 2.8.

multiple parameters: see paragraph 2.7.2.

numerical constraints: see paragraph 2.7.4.

odes: ordinary differential equations system. See paragraph 2.8.

optimal step: used when calculating new values of parameters during the iterative process. See paragraphs 2.6.1 and 2.6.2.

regression function: f. See paragraph 2.1.

replications: see page 4.

response: see paragraph 2.1.

starting values of parameters: see paragraph 2.6.3.

stopping criterion: $\mathcal{Q}_n(\theta, \beta)$. See page 13.

variance function: v. See paragraph 2.1.

weights: w. See paragraph 2.1.

3.2 Mathematical notations

$AsVar_{\widehat{\theta},\widehat{\beta}}$: asymptotic variance of the estimator of (θ, β). It is a $(p+q) \times (p+q)$ matrix. See page 9.

β: q vector of parameters defining the variance function v. See paragraph 2.1.

- β_0: starting value of β, see paragraph 2.6.3.
- $\widehat{\beta}$: estimation of β.
- $\widehat{\beta}_e$: estimation of β at step e, when (θ, β) are estimated alternatively. See page 9.
- β_{it}: current value of β at iteration it. See paragraph 2.6.

$B(\theta, \beta)$: $nh \times (p+q)$ matrix describing the estimating equations. See page 7

$\mathcal{C}_n$: *fitting criterion*. See page 12.

c: number of curves. See paragraph 2.7.2.

D: $nh \times (p+q)$ matrix of derivatives of η with respect to (θ, β). See page 8.

δ_{it}: *direction* at iteration it. See paragraph 2.6.

e: step number when (θ, β) are estimated alternatively. See page 9.

η: used in the definition of the estimating equations. $\eta = E(Z(Y))$. See page 7.

$\dot\eta_\theta$: $nh \times p$ matrix of derivatives of η with respect to θ. See page 8.

$\dot\eta_\beta$: $nh \times q$ matrix of derivatives of η with respect to β. See page 8.

f: regression function. f_l is the value of $f(x_l, \theta)$. l varies from 1 to n. See paragraph 2.1.

$\tilde{f}$: $n \times 1$ vector with components $f_1, \dots f_n$. See page 7.

$\tilde{f}^2$: $n \times 1$ vector with components $f_1^2, \dots f_n^2$. See page 7.

$\Delta(\tilde{f})$: diagonal matrix with components $f_1, \dots f_n$. See page 7.

$\partial f_l / \partial \theta$: $p \times 1$ vector of derivatives of f_l with respect to θ. See page 4.

$\dot{F}_\theta$: $n \times p$ matrix of derivatives of f with respect to θ, calculated at $x_1, \dots x_n$. See page 7

F_ν: functions occurring in the *odes*, ν varies from 1 to N_{eq}. F_ν depends on t, the integration variable, on unknown parameters θ_{odes}, on initial values of the system, and possibly on independent variables. See paragraph 2.8.

$F_{\nu,0}$: initial conditions of the *odes*, when they are known. When they are unknown, see item θ_{ci} (paragraph 3.2).

$\dfrac{\partial F_\nu}{\partial \theta}$: vector of derivatives of F_ν with respect to the parameters θ_{odes} and possibly θ_{ci}.

$\dfrac{dF_\nu}{dt}$: derivative of F_ν with respect to t.

I_{max}: maximum number of iterations. See paragraph 2.6.3.

it: current iteration number. See paragraph 2.6.

J: number of T values for which the *odes* has to be integrated, in each value of x. See paragraph 2.8.

k: number of distinct values of the pairs (x_l, w_l). See page 4.

λ, λ_1, λ_2, λ_{max}: occur in Gauss-Marquardt algorithm. See paragraph 2.6.2.

μ_{3l}: third order moment of Y_l. See page 8.

$\tilde\mu_3$: vector with components $\mu_{31}, \dots \mu_{3n}$.

$\Delta(\tilde\mu_3)$: diagonal matrix with components $\mu_{31}, \dots \mu_{3n}$. See page 8.

$\widehat{\mu}_{3l}$: estimation of μ_3. See page 9.

μ_{4l}: fourth order moment of Y_l. See page 8.

$\tilde\mu_4$: vector with components $\mu_{41}, \dots \mu_{4n}$.

$\Delta(\tilde\mu_4)$: diagonal matrix with components $\mu_{41}, \dots \mu_{4n}$. See page 8.

$\widehat{\mu}_{4l}$: estimation of μ_4. See page 9.

m: number of independent variables. See paragraph 2.1.

N_e: number of steps when the parameters (θ, β) are estimated alternatively. See page 9.

N_{eq}: number of equations defining the *odes*. See paragraph 2.8.

N_c: occurs in the iterative estimation process. See item ω_c (paragraph 3.2), item λ_2 (paragraph 3.2) and paragraph 2.6.

N_{err}: occurs in the iterative estimation process. See item ω_{err} (paragraph 3.2) and paragraph 2.6.3.

n: number of observations. See paragraph 2.1.

n_i: number of replications of Y, when $x = x_i$, $w = w_i$. See page 4.

nh: dimension of Z. See page 7

ω_c: occurs in Gauss-Newton algorithm. See paragraph 2.6.1.

ω_{err}: occurs in the iterative estimation proccess. See paragraph 2.6.3.

ω_{it}: *optimal step* at the end of iteration it. See paragraphs 2.6.1 and 2.6.2.

p: number of parameters occurring in the definition of f (dimension of θ). See paragraph 2.1.

p_{basic}: see paragraph 2.7.1.

p_{mult}: see paragraph 2.7.2.

p_{dist}: see paragraph 2.7.3.

p_{active}: see paragraph 2.7.4.

p_{odes}: see paragraph 2.8.

ϕ: gives the definition of f when f depends on the solution of an *odes*. See paragraph 2.8.

q: number of parameters occurring in the dimension of v and not in the definition of f (dimension de β). See paragraph 2.1.

q_{basic}: see paragraph 2.7.1.

q_{mult}: see paragraph 2.7.2.

q_{dist}: see paragraph 2.7.3.

q_{active}: see paragraph 2.7.4.

$\mathcal{Q}_n$: *fitting criterion* when the estimator is directly defined by *estimating equations*. See page 12, or see page 13 or item *stopping criterion* (paragraph 3.1).

$\mathcal{Q}_{stop}$: used to stop the iterative estimation process. See page 13 or item *stopping criterion* (paragraph 3.1).

r_l: error associated with Y_l. l varies from 1 to n. See paragraph 2.1.

$\widehat{r}_l$: residual $Y_l - f(x_l, \widehat{\theta})$. l varies from 1 to n. See page 9.

$\mathcal{R}_n$: sum of squared discrepancies between $(Y_l - f_l)^2$ and V_l divided by n. See page 6.

$\mathcal{S}_n$: residual sum of squares weighted by w_l/v_l, divided by n. See page 5.

s_i^2: empirical variance calculated using replications. See page 6.

$\Delta(\tilde{s}^2)$ diagonal matrix with components $s_1^2, \ldots s_n^2$. See page 8.

σ^2: parameter occurring in the description of the variance of Y_l. See paragraph 2.1.

$\widehat{\sigma}^2$: estimation of σ^2. See paragraph 2.3.5.

t: integration variable of the *odes*. It occurs in the description of f. See paragraph 2.8.

t_0: initial value of t. See paragraph 2.8.

T: $n \times J$ matrix with components the values of t. See paragraph 2.8.

$\mathcal{T}_n$: sum of r_l^2, divided by n. See page 5.

θ: $p \times 1$ vector of parameters occurring in the definition of f and possibly v. See paragraph 2.1.

θ_0: starting value of θ (see paragraph 2.6.3).

$\widehat{\theta}$: estimation of θ.

$\widehat{\theta}_e$: estimation of θ at step e, when the parameters (θ, β) are estimated alternatively. See page 9.

θ_{it}: current value of θ at iteration it. See paragraph 2.6.

θ_{odes}: set of parameters occurring in *odes*. See paragraph 2.8.

θ_{ci}: initial conditions of the *odes*. See paragraph 2.8.

θ_{ϕ}: set of parameters occurring in f but not in the *odes*. See paragraph 2.8.

$\mathcal{U}_n$: sum of squared errors, weighted by the empirical variances (calculated with replications), divided by n. See page 6.

$-2\log \mathcal{V}_n/n$: value of $-2\log$ of likelihood divided by n. See page 5.

$\dot{V}_\beta$: $n \times q$ matrix of derivatives of V with respect to β, calculated at $x_1, \ldots x_n$. See page 7.

$\dot{V}_\theta$: $n \times p$ matrix of derivatives of V with respect to θ, calculated at $x_1, \ldots x_n$. See page 7.

V_l: variance of Y_l: $V_l = \sigma^2 v(x_l, \theta, \beta)/w_l$. l varies from 1 to n. See paragraph 2.1.

$\tilde{V}$: $n \times 1$ vector with components V_1, V_n. See page 7.

$\Delta(\tilde{V})$: diagonal matrix with components $V_1, \ldots V_n$. See page 7.

$\partial V_l/\partial\beta$ **$q \times 1$:** vector of derivatives of V_l with respect to β. See page 4.

$\partial V_l/\partial\theta$**:** $p \times 1$ vector of derivatives of V_l with respect to θ. See page 4.

v**:** occurs in the description of the variance of Y and is called the *variance function*. v_l is the value of $v(x_l, \theta, \beta)$. l varies from 1 to n. See paragraph 2.1.

Var_Z**:** $nh \times nh$ covariance matrix of Z. See page 8.

W**:** $(p+q) \times (p+q)$ matrix defined by $W = B^T D/n$. See page 8.

w**:** *weighting variable*, occurs in the variance of Y. Its components are w_l, l varies from 1 to n. See paragraph 2.1.

X**:** $n \times m$ matrix of independent variables.

x_l**:** a $1 \times m$ row of X. See paragraph 2.1.

Y**:** observations. Y_l corresponds to the value x_l of the independent variables and verifies: $Y_l = f(x_l, \theta) + V_l^{1/2}\varepsilon_l$. l varies from 1 to n. See paragraph 2.1.

In case of an experimental design with replications, Y_{ij} corresponds to the value x_i of the independent variables. i varies from 1 to k, j from 1 to n_i. See page 4.

$\tilde{Y}$**:** $n\times$ vector of the observations Y_l. See page 7.

$\tilde{Y}^2$**:** $n \times 1$ vector with components $Y_1^2, \ldots Y_n^2$. See page 7.

Z**:** $nh \times 1$ vector which describes the estimating equations. See paragraph 2.4.

3.3 Notations used by nls2

The names used in the software cannot include mathematical symbols. Here is the correspondance between the notations used in it or in its on-line help-files and the mathematical notations of the paragraph 3.1.

algorithm: see paragraph 2.6

as.var: $AsVar_{\widehat{\theta},\widehat{\beta}}$ in dimension: number of *multiple parameters.*

B: $B(\theta, \beta)$.

B.varZ.B: $\frac{1}{n} B^T Var_Z B$ occurring in equation 2.2.

beta: $\widehat{\beta}$ in dimension: *multiple parameters.*

beta.start: β_0 in dimension: *multiple parameters.*

cond.start: $F_{\nu,0}$, $\nu = 1, \ldots N_{eq}$.

D: D.

d.resp: $\dfrac{\partial f_i}{\partial \theta}$, calculated at $(\widehat{\theta}, \widehat{\beta})$. The derivative are taken with respect to *multiple parameters.*

d.beta.vari: $\dfrac{\partial V_i}{\partial \beta}$, calculated at $(\widehat{\theta}, \widehat{\beta})$. The derivative are taken with respect to *multiple parameters.*

d.FOdes: $\dfrac{\partial F_\nu}{\partial \theta}(T_{l,\ell}, x_l, \widehat{\theta}_{odes})$ where the derivatives are taken with respect to θ_{odes}, or $\dfrac{\partial F_\nu}{\partial \theta}(T_{l,\ell}, x_l, \widehat{\theta}_{odes}, \widehat{\theta}_{ci})$ where the derivatives are taken with respect to $(\theta_{odes}, \theta_{ci})$. $\widehat{\theta}_{ci}$ are the estimated initial values of the system, if these values have been estimated. $\nu = 1, \ldots N_{eq}$, $\ell = 1, \ldots J$.

d.theta.vari: $\dfrac{\partial V_i}{\partial \theta}$, calculated at $(\widehat{\theta}, \widehat{\beta})$. The derivative are taken with respect to *multiple parameters.*

deriv.fct: $\dfrac{\partial f_i}{\partial \theta}$, $\dfrac{\partial V_i}{\partial \theta}$, $\dfrac{\partial V_i}{\partial \beta}$, calculated at the current value of (θ, β) during the iterative process, for $i = 1, \ldots k$. The derivatives are taken with respect to *active parameters.*

direction: $\delta_i t$.

Eta: η.

eq.beta: see paragraph 2.7.4.

eq.theta: see paragraph 2.7.4.

eqp.beta: see paragraph 2.7.3.

eqp.theta: see paragraph 2.7.3.

est.eq: B, D and η.

estim: current values of θ and β if any, during the iterative process. Same dimension as *active parameters.*

FOdes: $F_\nu(T_{l,\ell}, x_l, \widehat{\theta}_{odes})$ or $F_\nu(T_{l,\ell}, x_l, \widehat{\theta}_{odes}, \widehat{\theta}_{ci})$ if the initial values of the system are estimated, for $\nu = 1, \ldots N_{eq}$, $\ell = 1, \ldots J$.

f: f_i for $i = 1, \ldots k$

fitted: f_i and V_i for $i = 1, \ldots k$

gamf: second level parameters occurring in f (parameters that will not be estimated but fixed to a known value).

gamv: second level parameters occurring in v (parameters that will not be estimated but fixed to a known value).

inf.beta: see paragraph 2.7.4.

inf.theta: see paragraph 2.7.4.

integ.values: T.

iter: it.

lambda.c1: λ_1.

lambda.c2: λ_2

lambda.start: Gauss-Marquardt parameter at the first iteration. See paragraph 2.6.2.

loglik: $-2 \log \mathcal{V}_n/n$.

max.err.c1: N_{err}.

max.err.c2: N_c.

max.iters: I_{max}.

max.lambda: λ_{max}

max.stop.crit: $\mathcal{Q}_{stop}$.

method: see paragraph 2.2.

mu3, mu4: vectors with components μ_{3i} and μ_{4i}, for $i = 1, \dots k$.

mu.type: method for calculating the moments. See paragraph 2.5.

nb.iters: value of it at the end of the iterative procedure.

nb.steps: N_e

nb.theta.odes: p_{odes}.

norm: $\mathcal{Q}_n$ calculated at $(\widehat{\theta}, \widehat{\beta})$.

num.res: δ_{it}, ω_{it} and λ.

odes: F_ν, $\nu = 1, \dots N_{eq}$

omega: ω_{it}

omega.c1: ω_{err}.

omega.c2: ω_c.

phi: ϕ.

replications: n_i, $i = 1, \dots k$.

response: $f(x_i, \widehat{\theta})$, $i = 1, \dots k$

residuals: $\widehat{r}_l$, $l = 1, \dots n$.

rss: $n\mathcal{S}_n$ calculated at $(\widehat{\theta}, \widehat{\beta})$.

rss.unweighted: $n\mathcal{T}_n$ calculated in $(\widehat{\theta}, \widehat{\beta})$.

S2: s_i^2, $i = 1, \dots k$.

s.residuals: $\widehat{V}_l^{-1/2}\widehat{r}_l$, where $\widehat{V}_l = \widehat{\sigma}^2 v(x_l, \widehat{\theta}, \widehat{\beta})/w_l$, for $l = 1, \dots n$.

sigma2: value of σ^2.

sigma2.type: see paragraph 2.3.5.

stat.crit: $\mathcal{C}_n$.

start: t_0.

step: e.

stop.crit: $\mathcal{Q}_n$.

sup.beta: see paragraph 2.7.4.

sup.theta: see paragraph 2.7.4.

theta: $\widehat{\theta}$. Dimension: *multiple parameters*.

theta.start: θ_0. Dimension: *multiple parameters*.

v: v_i for $i = 1, \dots k$

vari.type: method for calculating the variance. See paragraph 2.2.

variance: $\widehat{\sigma}^2 v(x_i, \widehat{\theta}, \widehat{\beta})$ for $i = 1, \dots k$.

W: W.

weights: w.

Y1: $\dfrac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$, for $i = 1, \dots k$.

Y2: $\dfrac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}^2$, for $i = 1, \dots k$.

Z: Z.

Chapter 4

Methods for calibration

This chapter describes the numerical steps of the calibration function. See [7] for a complete description of the statistical background.

4.1 The problem

In calibration problems, we want to determine an unknown value z_0 of the *independent variable* — only one *independent variable* is possible in this context — from a measure Z of the *response*. m replicates $Z_1, Z_2, \ldots Z_m$ corresponding to the same value z_0 may be observed.

Function **calib.nls2** calculates confidence intervals for z_0 within bounds l_1, l_2.

4.2 Prerequisites

1. The regression parameters must have been previously estimated with a complete set of data (X, Y) denoted here by *standard data*.

2. The inverse of the regression function, denoted by f^{-1}, must be described by the user, either through a program or by means of a formal syntax.

3. It is assumed that the variances of Z and Y have the same structure.
   Examples:
   - $var(Y_l) = \sigma^2$ implies $var(Z) = \sigma^2$,
   - $var(Y_l) = \sigma^2 * v(x_l, \theta, \beta)$ implies $var(Z) = \sigma^2 * v(z_0, \theta, \beta)$.

4. The variance must be constant or the estimation method be a Maximum Likelihood method.

4.3 The calibration confidence sets

Two confidence sets are calculated.
- The first one, called S, is by construction an interval. It is based on the difference

between Z (or the mean of its m values, when several values of the response are observed) and the least squares estimator of z_0.

S is calculated only when the variance is constant.

- The second one, called R, may not be an interval. It is based on the signed root of the likelihood ratio.

Note: a comparison of the calibration methods can be found in [7].

4.4 The confidence interval S

The following quantities are calculated:

1. the residual sums of squares $\widehat{S}$:

$$\widehat{S} = \sum_{i=1}^{n}(Y_i - f(x_i, \widehat{\theta}))^2 + \sum_{j=1}^{m}(Z_j - \bar{Z})^2 \tag{4.1}$$

   where
   $\bar{Z}$ is the mean of the m values of Z,
   n is the total number of observations of the *standard data*,
   $\widehat{\theta}$ is the estimation of θ calculated with the *standard data*.

2. the estimator of z_0, $\widehat{z}$:
   Suppose f is an increasing function.
   - if $\bar{Z}$ belongs to the interval $\left]f(\ell_1, \widehat{\theta}), f(\ell_2, \widehat{\theta})\right[$, then $\widehat{z} = f^{-1}(\bar{Z}, \widehat{\theta})$,
   - if $\bar{Z}$ is less than $f(l_1, \widehat{\theta})$, then $\widehat{z} = l_1$,
   - if $\bar{Z}$ is greater than $f(l_2, \widehat{\theta})$, then $\widehat{z} = l_2$.

   If f is decreasing, $\widehat{z}$ is determined by a similar procedure.

3. the variance term A:

$$A = \sqrt{n+m} * \frac{1}{\sqrt{\widehat{S}} * \sqrt{1/m + (\frac{\partial f(\widehat{z},\widehat{\theta})}{\partial \theta} * \frac{AsVar_{\widehat{\theta}}}{\widehat{\sigma}^2} * \frac{\partial f(\widehat{z},\widehat{\theta})}{\partial \theta}^t)}}$$

   where $\widehat{\sigma}^2$ is the estimation of σ^2 calculated with the *standard data*.

The values z in the confidence interval verify:

$$u_\alpha \leq S(z) \leq u_{1-\alpha} \tag{4.2}$$

where

$$S(z) = A * (\bar{Z} - f(z, \widehat{\theta}))$$

α is a predetermined confidence level, $(0 \leq \alpha \leq 1)$, and u_α is defined by $\Phi(u_\alpha) = \alpha$, Φ being the distribution function of a standard normal random variable.

Therefore, the S interval is:

$$\left[f^{-1}(\bar{Z} - \frac{u_\alpha}{A}, \widehat{\theta}), f^{-1}(\bar{Z} - \frac{u_{1-\alpha}}{A}, \widehat{\theta}\right]$$

4.4.1 A bootstrap version of S

The bootstrap method replaces the normal bounds u_α and $u_{1-\alpha}$ (see 4.2) by bootstrap bounds. You can obtain them by using function **bootstrap.nls2** with argument **method** equal to **calib**.

4.5 The confidence set R

The confidence set R is based on a statistic which is calculated at several points z of a regularly spaced grid defined in $[l_1, l_2]$. If the value $R(z)$ satisfies the condition:

$$u_\alpha \leq R(z) \leq u_{1-\alpha} \tag{4.3}$$

then z is a point of the confidence set.
(u_α is defined in paragraph 4.4).

4.5.1 The numerical procedure

When the variance is constant

For each value z of the grid:

1. The point (z, Z) is added to the *standard data*, with its m replications, if any.
2. An estimation is run.
   The sum of squares is $\widetilde{S} = \sum_{i=1}^{m+n} \widetilde{r}_i^2$, where $\widetilde{r}$ denotes the residuals.
3. $R(z)$ is given by:

$$R(z) = sign(\bar{Z} - f(z, \widehat{\theta})) * \sqrt{(m+n) * log(\frac{\widetilde{S}}{\widehat{S}})}$$

   $\widehat{S}$ is defined in equation (4.1), paragraph 4.4.

When the variance is not constant

1. The point ($\widehat{z}$, Z) is added to the *standard data*, with its m replications, if any.
   ($\widehat{z}$ is defined at point 2, paragraph 4.4).
2. An estimation is run, the unknown abscissa been estimated along with the regression parameters.
   (The abscissa of the added point is reset at each iteration).

New values are so calculated:
- the estimation of the regression parameters $\widehat{\theta}$,
- the estimation of the unknown abscissa $\widehat{z}$,
- the statistical criterion $\widehat{L} = -2log(likelihood)/(n+m)$.

3. Then, for each value z of the grid:

   (a) The point $(z,\ Z)$ is added to the *standard data*, with its m replications, if any.

   (b) An estimation is run, where the regression parameters are estimated only. A new value of the statistical criterion $\widetilde{L}$ is so calculated.

   (c) $R(z)$ is given by:

$$R(z) = sign(\bar{Z} - f(z, \widehat{\theta})) * \sqrt{-(m+n) * (\widehat{L} - \widetilde{L})}$$

For a more accurate determination of R

The grid of research may be split recursively: for each consecutive points z_i, z_{i+1} so that z_i belongs to R and z_{i+1} does not (or vice-versa), the interval $[z_i, z_{i+1}]$ is split again and determines a new grid of reasearch.

4.6 List of notations used in the calibration

These notations complete or replace the notations of chapter 3.

4.6.1 Key words

standard data: data set used to estimate the parameters of the model. See paragraph 4.2.

4.6.2 Mathematical notations

l_1, l_2: lower and upper bounds of the interval where z_0 should lie. Confidence intervals are researched within these bounds. See paragraph 4.1.

α: confidence level.

u_α is defined by $\Phi(u_\alpha) = \alpha$, Φ being the distribution function of a standard normal random variable. See paragraph 4.4, equation (4.2) and paragraph 4.5, equation (4.3).

m: number of replications of z_0, i.e number of values of Z. See paragraph 4.1.

n: total number of observations of the *standard data* (replications included). See equation 4.1.

$\widehat{L}$: statistical criterion calculated when the unknown abscissa is estimated along with the parameters of the regression model. See paragraph 4.5.1, variance not constant.

$\widehat{S}$: residual sums of squares calculated with the $m+n$ observations and $(\widehat{\theta}, \widehat{\sigma^2})$. See equation (4.1), paragraph 4.4.

$\widetilde{S}$: residual sums of squares calculated with the $m+n$ observations and the current estimation of the parameters. See paragraph 4.5.1, item 1.

$\widehat{\theta}$: estimations of the parameters calculated with the *standard data* (see equation 4.1) or calculated along with the unknown abscissa (see paragraph 4.5.1,variance not constant).

$\widehat{\sigma^2}$: estimation of σ^2 calculated with the *standard data*. See paragraph 4.4, point 3.

Z: observed response corresponding to the unknown value of the independent variable z_0. m replications $Z_i, i = 1 \dots m$ may be observed. See paragraph 4.1.

$\bar{Z}$: mean of the m replications of Z. See paragraph 4.4.

z_0: value of the independent variable to be estimated. See paragraph 4.1.

$\widehat{z}$: estimator of z_0. See paragraph 4.4, item 2, and paragraph 4.5.1.

4.6.3 Notations used by nls2

Here is the correspondance between the notations used in the software and the mathematical notations used in the paragraph 4.6.2.

Inputs

conf.bounds: quantiles of the $S(z)$ and $R(z)$ distributions. When they are set, they replace conf.level.

conf.level: required confidence level, i.e α. See paragraph 4.4, equation (4.2) and paragraph 4.5, equation (4.3).

ord: observed values of the response Z. See paragraph 4.1

R.grid: minimum number of points of the research grid for R. See paragraph 4.5

R.nsplit: number of times a new research grid is built when a break is encountered during the determination of R. See *Note* in paragraph 4.5.1.

x.bounds: lower and upper bounds of the research interval, i.e l_1 and l_2. See paragraph 4.1.

Outputs

S.conf.int, R.conf.int: confidence limits of the intervals S and R respectively. R.conf.int exists only when R is an interval.

R.conf.set: points of the research grid that satisfy (4.3). R.conf.set replaces R.conf.int when R is not an interval.

R.values: values of $R(z)$ at the points z of the research grid. See paragraph 4.5.1, item 2.

R.x: points of the research grid, i.e the different values of z in paragraph 4.5.1.

x: estimator of z_0, i.e $\widehat{z}$. See paragraph 4.4, item 2.

Bibliography

[1] O. Bunke. Estimating the accuracy of estimators under regression models. Technical report, Humboldt University, Berlin, 1990.

[2] N.L. Carr. Kinetics of catalytic isomerization of n-pentane. *Industrial and Engineering Chemistry*, 52:391–396, 1960.

[3] D.G. Watts D.M. Bates. *Nonlinear regression analysis and its applications.* Wiley, New York, 1988.

[4] J.P. Renard F. Legal, P. Gasqui. Differential osmotic behaviour of mammalian oocytes before and after maturation. *Cryobiology.*

[5] D.J. Finney. *Statistical Method in Biological Assay.* Griffin, London, 1978.

[6] C.J. Wild G.A.F. Seber. *Nonlinear Regression.* Wiley, New-York, 1989.

[7] M.A. Gruet and E. Jolivet. Calibration with a nonlinear standard curve: how to do it.? *Computational Statistics*, 4:249–276, 1994.

[8] A.C. Hindmarsh. *A systematized collection of ODE solvers, Vol. 1.* Scientific Computing, North Holland, 1983.

[9] K.M.S. Humak. *Statistische Methoden der Modellbildung*, volume 2. Nichtlineare Regression, robuste Verfahren in linearen Modellen. Modelle mit Fehlern in den variablen. Akademie-Verlag, Berlin, 1983.

[10] C. Millier J.D. Lebreton. *Modèles Dynamiques Déterministes en Biologie.* Masson, Paris, 1982.

[11] L. Petzold. Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM J. Sci. Stat. Comput.*, 4:??–??, 1983.

[12] D.A. Ratkowsky. *Nonlinear Regression Modeling.* M. Dekker, New-York, 1983.

[13] D.A. Ratkowsky. *Handbook of Nonlinear Regression Models.* Dekker, New-York, 1989.

[14] A. Messéan S. Huet, E. Jolivet. *La régression non-linéaire: méthodes et applications à la biologie.* INRA, Paris, 1991.

[15] E. Jolivet S. Huet and A. Messéan. Some simulations results about confidence intervals and bootstrap methods in nonlinear regression. *Statistics*, 21:369–432, 1989.

[16] E. Jolivet S. Huet, A.M. Gruet and A. Bouvier. *Statistical tools for nonlinear regression: a pratical guide.* Springer-Verlag, New-York, 1996.

[17] J. Laporte S. Huet and J.F. Vautherot. Statistical methods for the comparison of antibody levels in serums assayed by enzyme linked immuno sorbent assay. *Biométrie–Praximétrie*, 28:61–80, 1987.

[18] C.F.J Wu. Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *The Annals of Statistics*, 14:1291–1380, 1986.