The R package nlstools: a toolbox for nonlinear regression

Florent Baty
Sandrine Charles
Jean-Pierre Flandrois
Marie-Laure Delignette-Muller

10/07/2009
Introduction

- Increasing interest for nonlinear regression
  - Chemistry
  - Agricultural science
  - Pharmacology
  - Microbiology

- Diversity of tools available in R (Ritz & Streibig, 2008)
Many users still reluctant in using nonlinear regression

- Complexity of the least-squares minimization procedure algorithms
- Definition of starting values
- Validity of the error model
- Confidence region estimation

**nlstools**

- A toolbox which helps the fit of gaussian nonlinear models and assess its quality of fit

\[ y_i = f(x_i, \theta) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma) \] (1)

- Available on CRAN
Overview of nlstools

- Main features of nlstools

Help for the fit

Extended summary

Validity of the error model

Resampling
  - Bootstrapping
  - Jackknifing

Parameters’ confidence regions

Collection of models and data sets

florent baty (kssg)
r package nlstools

Help of fit: preview

- Estimation of starting values by iterative "manual" fitting

> preview (formula, data, start, variable = 1)
Help of fit: preview

- Models with non-meaningful parameters (Ratkowsky)

\[ \mu_{\text{max}} = b(pH - pH_{\text{min}})(1 - \exp(c(pH - pH_{\text{max}})))^2 \]  

(2)
Central function

```r
> nls (formula, data, start, ...)
```

`nlstools` requires objects of class `nls`
Fit summary: plotfit, overview

- Extended summary of fit

```r
> plotfit (nls)
> overview (nls)
```

- Formula
- Parameters estimates and standard error
- Residual sum of squares
- Asymptotic confidence intervals
- Parameters correlation matrix
Error model: `nlsResiduals`, `test.nlsResiduals`

- Residuals analysis: validation of the error model

```r
> plot(nlsResiduals(nls))
> test.nlsResiduals(nlsResiduals(nls))
```

- Normality test:
  - Shapiro-Wilk

- Test of independence:
  - Runs test
Advanced features: \texttt{nlsConfRegions}

- \textbf{Projections} of Beale 95\% confidence region (Beale, JRSS, 1960)

\[
\text{RSS}(\theta) < \text{RSS}_{\text{min}}[1 + \frac{p}{n - p} F_{1-\alpha}(p, n - p)]
\]  \hspace{1cm} (3)

\[> \text{plot(nlsConfRegions(nls))}\]
Advanced features: `nlsContourRSS`

- Contour of the RSS (+ sections of Beale 95% confidence region)

```r
> plot(nlsContourRSS(nls))
```

Florent Baty (KSSG)
Advanced features: \texttt{nlsContourRSS}

- Problems of ill-conditioning in parameter estimation (Ratkowsky model)

\texttt{> plot(nlsContourRSS(nls))}
Advanced features: `nlsConfRegions` vs. `nlsContourRSS`

- Comparison of both representations of the confidence region: **Projections** vs. **sections** of the Beale region

- Both representations equivalent in models with 2 parameters
- Different when number of parameters $\geq 2$
- Projections more suitable to assess global correlations between parameters
Advanced features: \texttt{nlsJack}

- Jackknifing

\begin{verbatim}
> plot(nlsJack(nls))
> summary(nlsJack(nls))
\end{verbatim}

- Jackknife estimate
- Confidence interval
- Influential observations
Advanced features: nlsBoot

- Nonparametric bootstrapping of mean centered residuals
- Bootstrap estimates and confidence intervals

```r
> plot(nlsBoot(nls))
```

```r
> plot(nlsBoot(nls), type="boxplot")
```
**Summary**

- **nlstools** includes a set of tools which extend the functionalities of nonlinear regression in R.
- A particular attention was paid in the development of:
  - representations of confidence regions
  - resampling techniques (jackknifing and bootstrapping)
- **nlstools** is particularly helpful
  - to fit models with non-meaningful parameters
  - to detect problematic models (overparameterized, highly correlated parameters, ill-conditioning estimation in parameter estimation, ...)
Conclusion

- **nlstools**: user-friendly interface of a set of basic and more advanced diagnostic functions in the framework of gaussian nonlinear regression
- Available on CRAN, including a didactic vignette
  - http://cran.r-project.org/web/packages/nlstools
- Some improvements:
  - Hypervalidation
  - Better flexibility of graphics