Profiling the parameters of models with linear predictors

by Ioannis Kosmidis
Department of Statistics, University of Warwick
Coventry, CV4 7AL, UK

March 27, 2008

Profiles of the likelihood can be used for the construction of confidence intervals for parameters, as well as to assess features of the likelihood surface such as local maxima, asymptotes, etc., which can affect the performance of asymptotic procedures. The `profile` methods of the **R** language (**stats** and **MASS** packages) can be used for profiling the likelihood function for several classes of fitted objects, such as `glm` and `polr`. However, the methods are limited to cases where the profiles are almost quadratic in shape and can fail, for example, in cases where the profiles have an asymptote.

Furthermore, often the likelihood is replaced by an alternative objective for either the improvement of the properties of the estimator, or for computational efficiency when the likelihood has a complicated form (see, for example, Firth (1993) for a maximum penalized likelihood approach to bias-reduction, and Lindsay (1988) for composite likelihood methods, respectively). Alternatively, estimation might be performed by using a set of estimating equations which do not necessarily correspond to a unique objective to be optimized, as in quasi-likelihood estimation (Wedderburn, 1974; McCullagh, 1983) and in generalized estimating equations for models for clustered and longitudinal data (Liang & Zeger, 1986). In all of the above cases, the construction of confidence intervals can be done using the profiles of appropriate objective functions in the same way as the likelihood profiles. For example, in the case of bias-reduction in logistic regression via maximum penalized likelihood, Heinze & Schemper (2002) suggest to use the profiles of the penalized likelihood, and when estimation is via a set of estimating equations Lindsay & Qu (2003) suggest the use of profiles of appropriate quadratic score functions.

In this presentation we introduce the **profileModel** package, which generalizes the capabilities of the current `profile` methods to arbitrary, user-specified objectives and, also, covers a variety of current and potentially future implementations of fitting procedures that relate to models with *linear predictors*. We give examples of how the package can be used to calculate, evaluate and plot the profiles of the objectives, as well as to construct profile-based confidence intervals. The presentation focuses on the following:

- **Generality of application:** The **profileModel** package has been designed to support classes of fitted objects with linear predictors that are constructed according to the specifications given by Chambers & Hastie (1991, Chapter 2). Such generality of application stems from the appropriate use of generic **R** methods such as `model.frame`, `model.matrix` and `formula`.

- **Embedding:** The developers of current and new fitting procedures as well as the end-users can have direct access to profiling capabilities. The only requirement is authoring a simple function that calculates the value of the appropriate — for their specific application — objective to be profiled.

- **Computational stability:** All the facilities have been developed with computational stability in mind, in order to provide an alternative which improves and extends the capabilities of already available `profile` methods.

References

Chambers, J. M. & Hastie, T. (1991). *Statistical Models in S.* Chapman & Hall.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika* **80**, 27–38.

Heinze, G. & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine* **21**, 2409–2419.

Liang, K.-Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Lindsay, B. G. (1988). Composite likelihood methods. In *Statistical Inference from Stochastic Processes*, Ed. N. U. Prabhu, pp. 221–239. American Mathematical Society.

Lindsay, B. G. & Qu, A. (2003). Inference functions and quadratic score tests. *Statistical Science* **18**, 394–410.

McCullagh, P. (1983). Quasi-likelihood functions. *The Annals of Statistics* **11**, 59–67.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 439–447.