maxLik: A Package for Maximum Likelihood Estimation in R

Ott Toomet[1,2] **and Arne Henningsen**[3,4,*]

1. Department of Economics, University of Tartu (Estonia)
2. Department of Economics, Aarhus School of Business, University of Aarhus (Denmark)
3. Department of Agricultural Economics, University of Kiel (Germany)
4. Institute of Food and Resource Economics, University of Copenhagen (Denmark)

* Contact author: arne.henningsen@gmail.com

Keywords: Maximum Likelihood, Optimisation

The **maxLik** package provides convenient tools for maximum likelihood (ML) estimations in the statistical software environment R. This package is available from CRAN (`http://cran.r-project.org/package=maxLik`), R-Forge (`http://r-forge.r-project.org/projects/maxlik/`), and its homepage (`http://www.maxLik.org/`).

The most important tool for a user of the **maxLik** package is probably the `maxLik` function. It is a wrapper function that delegates the maximum likelihood estimation to the selected optimisation routine. Five optimisation methods are currently available (names of the corresponding functions in parenthesis): Newton-Raphson (`maxNR`), Berndt-Hall-Hall-Hausman (`maxBHHH`), Broyden-Fletcher-Goldfarb-Shanno (`maxBFGS`), Nelder-Mead (`maxNM`), and simulated-annealing (`maxSANN`). While the actual optimisation in `maxBFGS`, `maxNM`, and `maxSANN` is done by `optim`, the Newton-Raphson algorithm is implemented in the function `maxNR` itself. The actual optimisation in `maxBHHH` is done by `maxNR`.

The first argument of `maxLik` (`loglik`) is mandatory and specifies the log-likelihood function. Its first argument must be the vector of the parameters to be estimated and it must return either a single log-likelihood value or a numeric vector where each component is the log-likelihood value corresponding to an individual observations. The second and third argument (`grad` and `hess`) are optional and can be used to specify functions that return the gradients and the Hessian of the objective function, respectively. If these functions are not provided by the user, numerical gradients and Hessians are calculated if necessary. The fourth argument (`start`) is mandatory and must be used to specify a vector of starting values. Finally, the fifth argument (`method`) is optional and can be used to select the maximisation routine. It defaults to `"NR"`, but it can also be `"BHHH"`, `"BFGS"`, `"NM"`, or `"SANN"`. The `maxLik` wrapper capabilities are designed in a transparent way, so that the user can easily swap the methods without changing the arguments. The arguments not used by a particular optimisation method, such as `hess` for the Berndt-Hall-Hall-Hausman method, are ignored.

The **maxLik** package is implemented using S3 classes. The `maxLik` wrapper returns a list of class `"maxLik"`. Corresponding methods can handle the likelihood-specific properties of the estimate including the fact that inverse of the negative Hessian is the variance-covariance matrix of the estimated parameters. The most important methods for objects of class `"maxLik"` are: `summary` for returning (and printing) summary results, `coef` for extracting the estimated parameters, `vcov` for calculating the variance covariance matrix of the estimated parameters, `logLik` for extracting the log likelihood value, and `AIC` for calculating the Akaike information criterion.

Currently, the **maxLik** package is used for maximum likelihood estimations in three packages that are available on CRAN: **mlogit**, **sampleSelection**, and **truncreg**. On the useR! conference, we would like to demonstrate how to use the **maxLik** package for maximum likelihood estimations in R. Furthermore, we would like to highlight its advantages and features to encourage more users and package writers to use the **maxLik** package.