Package 'cAIC4'

February 19, 2015

Type Package

Title Conditional Akaike information criterion for lme4

Version 0.2

Date 2014-05-23

Author Benjamin Saefken and David Ruegamer, with contributions from Sonja Greven and Thomas Kneib

Maintainer Benjamin Saefken <bsaefke@uni-goettingen.de>

Depends lme4(>= 1.1-6), methods, Matrix

Suggests gamm4

Description Provides functions for the estimation of the conditional Akaike information in generalized mixed-effects models fitted with (g)lmer form lme4.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2014-08-12 11:48:10

R topics documented:

cAIC4-package *Conditional Akaike information criterion for lme4*

Description

Provides functions for the estimation of the conditional Akaike information in generalized mixed-effects models fitted with (g)lmer form lme4.

Details

Package:	cAIC
Type:	Package
Version:	0.2
Date:	2014-05-23
License:	GPL (>=2)

Author(s)

Benjamin Saefken, David Ruegamer, Thomas Kneib and Sonja Greven.

Maintainer: Benjamin Saefken <bsaefke@uni-goettingen.de>

References

Saefken, B., Kneib T., van Waveren C.-S. and Greven, S. (2014) A unifying approach to the estimation of the conditional Akaike information in generalized linear mixed models. Electronic Journal Statistics Vol. 8, 201-225.

Greven, S. and Kneib T. (2010) On the behaviour of marginal and conditional AIC in linear mixed models. Biometrika 97(4), 773-789.

Efron , B. (2004) The estimation of prediction error. J. Amer. Statist. Ass. 99(467), 619-632.

Examples

```
b <- lmer(Reaction ~ Days + (Days | Subject), sleepstudy)

cAIC(b)
```

cAIC *Conditional Akaike information for lmer & glmer*

Description

Estimates the conditional Akaike information for models that were fitted in lme4. This is possible for all distributions, i.e. `family` arguments, based on parametric conditional bootstrap. For the Gaussian distribution (from a `lmer` call) and the Poisson distribution analytical estimators for the degrees of freedom are available, based on Stein type formulas. Also the conditional Akaike information for generalized additive models based on a fit via the gamm4-package can be estimated.

Usage

```
cAIC(object, method = NULL, B = NULL, sigma.estimated = TRUE, analytic = TRUE)
```

Arguments

`object`	An object of class merMod either fitted by `lmer` or `glmer` of the lme4-package. Also objects returned form a gamm4 call are possible.
`method`	Either `"conditionalBootstrap"` for the estimation of the degrees of freedom with the help of conditional Bootstrap or `"steinian"` for analytical representations based on Stein type formulas. The default is `NULL`. In this case the method is choosen automatically based on the `family` argument of the `(g)lmer`-object. For `"gaussian"` and `"poisson"` this is the Steinian type estimator, for all others it is the conditional Bootstrap.
`B`	Number of Bootstrap replications. The default is `NULL`. Then B is the minimum of 100 and the length of the response vector.
`sigma.estimated`	If sigma is estimated. Only used for the analytical version of Gaussian responses.
`analytic`	FALSE if the numeric hessian of the (restricted) marginal log-likelihood from the lmer optimization procedure should be used. Otherwise (default) TRUE, i.e. use a analytical version that has to be computed. Only used for the analytical version of Gaussian responses.

Details

For `method = "steinian"` and an object of class `merMod` computed the analytic representation of the corrected conditional AIC in Greven and Kneib (2010). This is based on a the Stein formula and uses implicit differentiation to calculate the derivative of the random effects covariance parameters w.r.t. the data. The code is adapted form the one provided in the supplementary material of the paper by Greven and Kneib (2010). The supplied `merMod` model needs to be checked if a random effects covariance parameter has an optimum on the boundary, i.e. is zero. And if so the model needs to be refitted with the according random effect terms omitted. This is also done by the function and the refitted model is also returned. Notice that the `boundary.tol` argument in `lmerControl` has an impact on whether a parameter is estimated to lie on the boundary of the parameter space. For

estimated error variance an the degrees of freedom are increased by one. If this should not be done set `sigma.estimated = "FALSE"`.

If the object is of class `merMod` and has `family = "poisson"` there is also an analytic representation of the conditional AIC based on the Chen-Stein formula, see for instance Saefken et. al (2014). For the calculation the model needs to be refitted for each observed response variable minus the number of response variables that are exactly zero. The calculation therefore takes longer then for models with Gaussian responses. Due to the speed and stability of lme4 this is still possible, also for larger datasets.

If the model has Bernoulli distributed responses and `method = "steinian"`, `cAIC` calculates the degrees of freedom based on a proposed estimator by Efron (2004). This estimator is asymptotically unbiased if the estimated conditional mean is consistent. The calculation needs as many model refits as there are data points.

Another more general method for the estimation of the degrees of freedom is the conditional bootstrap. This is proposed in Efron (2004). For the B boostrap samples the degrees of freedom are estimated by

$$\frac{1}{B-1}\sum_{i=1}^{n}\theta_i(z_i)(z_i-\bar{z}),$$

where $\theta_i(z_i)$ is the i-th element of the estimated natural parameter.

Value

A list consisting of: 1. the conditional log likelihood, i.e. the log likelihood with the random effects as penalized parameters; 2. the estimated degrees of freedom; 3. a list element that is either `NULL` if no new model was fitted otherwise the new (reduced) model, see details; 4. a boolean variable indicating whether a new model was fitted or not; 5. the estimator of the conditional Akaike information, i.e. minus twice the log likelihood plus twice the degrees of freedom.

WARNINGS

Currently the cAIC can only be estimated for `family` equal to `"gaussian"`, `"poisson"` and `"binomial"`. Neither negative binomial nor gamma distributed responses are available.

Weighted Gaussian models are not yet implemented.

Author(s)

Benjamin Saefken <bsaefke@uni-goettingen.de>

References

Saefken, B., Kneib T., van Waveren C.-S. and Greven, S. (2014) A unifying approach to the estimation of the conditional Akaike information in generalized linear mixed models. Electronic Journal Statistics Vol. 8, 201-225.

Greven, S. and Kneib T. (2010) On the behaviour of marginal and conditional AIC in linear mixed models. Biometrika 97(4), 773-789.

Efron , B. (2004) The estimation of prediction error. J. Amer. Statist. Ass. 99(467), 619-632.

See Also

lme4-package, lmer, glmer

Examples

```
b <- lmer(Reaction ~ Days + (Days | Subject), sleepstudy)

cAIC(b)

b2 <- lmer(Reaction ~ (1 | Days) + (1 | Subject), sleepstudy, analytic = FALSE)

cAIC(b2)

b2ML <- lmer(Reaction ~ (1 + Days | Subject), sleepstudy, REML = FALSE)

cAIC(b2ML)
```

---

deleteZeroComponents *Delete random effect terms with zero variance*

Description

Is used in the cAIC function if `method = "steinian"` and `family = "gaussian"`. The function deletes all random effects terms from the call if corresponding variance parameter is estimated to zero and updates the model in merMod.

Usage

```
deleteZeroComponents(m)
```

Arguments

m	An object of class merMod fitted by lmer of the lme4-package.

Details

Uses the cnms slot of m and the relative covariance factors to rewrite the random effects part of the formula, reduced by those parameters that have an optimum on the boundary. This is necessary to obtain the true conditional corrected Akaike information. For the theoretical justification see Greven and Kneib (2010). The reduced model formula is then updated. The function deleteZero-Components is then called iteratively to check if in the updated model there are relative covariance factors parameters on the boundary.

Value

An updated object of class merMod

WARNINGS

For models called via gamm4 no automated update is available. Instead a warning with terms to omit from the model is returned.

Author(s)

Benjamin Saefken <bsaefke@uni-goettingen.de> \& David Ruegamer

References

Greven, S. and Kneib T. (2010) On the behaviour of marginal and conditional AIC in linear mixed models. Biometrika 97(4), 773-789.

See Also

lme4-package, lmer, getME

Examples

```
## Currently no data with variance equal to zero...
b <- lmer(Reaction ~ Days + (Days | Subject), sleepstudy)

deleteZeroComponents(b)
```

Index