Package 'mglm4twin'

August 1, 2018

Type Package

Title Multivariate Generalized linear models for twin data

Version 0.1.0

Date 2018-06-05

Author Wagner Hugo Bonat [aut, cre],
    Jacob v. B. Hjelmborg [ctb]

Maintainer Wagner Hugo Bonat <wbonat@ufpr.br>

Description Fitting multivariate generalized linear models for twin and family data.

Depends R (>= 3.2.1)

Suggests testthat, knitr, devtools

Imports stats, Matrix, assertthat, graphics, Rcpp (>= 0.12.16)

License GPL-3 | file LICENSE

LazyData TRUE

URL https://github.com/wbonat/mglm4twin

BugReports https://github.com/wbonat/mglm4twin/issues

Encoding UTF-8

VignetteBuilder knitr

RoxygenNote 6.0.1

LinkingTo Rcpp, RcppArmadillo

NeedsCompilation yes

R topics documented:

coef.mglm4twin *Model Coefficients*

Description

Extract model coefficients for objects of mglm4twin class.

Usage

```
## S3 method for class 'mglm4twin'
coef(object, std.error = FALSE, model, response = NULL,
  ...)
```

Arguments

object	an object of mglm4twin class.
std.error	logical. If TRUE returns the standard errors for the estimates. Default is FALSE.
model	Type of the fitted model. Options are E, AE, CE, ACE and ADE.
response	For which response regression coefficients are extracted. Default = NULL. It means all estimates are returned including dispersion estimates for all responses.
...	additional arguments affecting the summary produced. Note that there is no extra options for mglm4twin object class.

Value

A data.frame with parameters names, estimates, response variable number and parameters type.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

<pre>
  fit_mglm   <i>Chaser and Reciprocal Likelihood algorithms</i>
</pre>

Description

This function implements the two main algorithms used for fitting multivariate generalized linear models to twin data, i.e. The chaser and the reciprocal likelihood algorithms.

Usage

```
fit_mglm(list_initial, list_link, list_variance,
         list_X, list_Z, list_offset, list_Ntrial, list_power_fixed,
         list_weights, y_vec, correct, max_iter, tol, method,
         tuning, verbose)
```

Arguments

list_initial	a list of initial values for regression and covariance parameters.
list_link	a list specifying the link function names. Options are: "logit", "probit", "cauchit", "cloglog", "loglog", "identity", "log", "sqrt", "1/mu^2" and "inverse". See mt_link_function for details. Default link = "identity".
list_variance	a list specifying the variance function names. Options are: "constant", "tweedie", "poisson_tweedie", "binomialP" and "binomialPQ". See mt_variance_function for details. Default variance = "constant".
list_X	a list of design matrices. See model.matrix for details.
list_Z	a list of knowm matrices to compose the matrix linear predictor.
list_offset	a list of offset values. Default NULL.
list_Ntrial	a list of number of trials, useful only when analysing binomial data. Default 1.
list_power_fixed	
	a list of logicals indicating if the power parameters should be estimated or not. Default power_fixed = TRUE.
list_weights	Additional weights for the quasi-score function. (Default = NULL).
y_vec	a vector of the stacked response variables.
correct	a logical indicating if the algorithm will use the correction term or not. Default correct = FALSE.
max_iter	maximum number of iterations. Default max_iter = 20.
tol	a numeric specyfing the tolerance. Default tol = 1e-04.
method	a string specyfing the method used to fit the models ("chaser" or "rc"). Default method = "chaser".
tuning	a numeric value in general close to zero for the rc method and close to 1 for the chaser method. This argument control the step-length. Default tuning = 1.
verbose	a logical if TRUE print the values of the covariance parameters used on each iteration. Default verbose = FALSE

Value

A list with estimated regression and covariance parameters. Details about the estimation procedures as iterations, sensitivity, variability are also provided. In general the users do not need to use this function directly. The `mglm4twin` provides GLM interface for fitting multivariate generalized linear models for twin data.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

Source

Bonat, W. H. and Jorgensen, B. (2016) Multivariate covariance generalized linear models. Journal of Royal Statistical Society - Series C 65:649–675.

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

See Also

mglm4twin, mt_matrix_linear_predictor, mt_link_function and mt_variance_function.

---

gof   *Measures of Goodness-of-Fit*

Description

Extract the pseudo Gaussian log-likelihood (plogLik), pseudo Akaike Information Criterion (pAIC), pseudo Kullback-Leibler Information Criterion (pKLIC) and pseudo Bayesian Information Criterion (pBIC) for objects of `mglm4twin` class.

Usage

```
gof(object)
```

Arguments

object   an object or a list of objects representing a model of `mglm4twin` class.

Value

Returns a data frame containing goodness-of-fit measures.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

Source

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

Wang, M. (2014). Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments. Advances in Statistics, 1(1)1–13.

See Also

plogLik, pAIC, pKLIC and pBIC.

| mglm4twin | *Fitting Multivariate Covariance Generalized Linear Models to Twin data* |

Description

The function mglm4twin is used for fitting multivariate generalized linear models to twin data. The models are specified by a set of lists giving a symbolic description of the linear and matrix linear predictors. The user can choose between a list of link and variance functions. The models are fitted using an estimating function approach, combining the quasi-score function for the regression parameters and the Pearson estimating function for the covariance parameters. For details see Bonat and Jorgensen (2016).

Usage

```
mglm4twin(linear_pred, matrix_pred, link, variance,
        offset, Ntrial, power_fixed, weights, data, control_initial,
        contrasts, control_algorithm)
```

Arguments

linear_pred	a list of formula see [formula](#) for details.
matrix_pred	a list of matrices, in general the output of mt_twin function.
link	a list of link functions names. Options are: "logit", "probit", "cauchit", "cloglog", "loglog", "identity", "log", "sqrt", "1/mu^2" and "inverse". See [mt_link_function](#) for details.
variance	a list of variance functions names. Options are: "constant", "tweedie", "poisson_tweedie", "binomialP" and "binomialPQ". See [mt_variance_function](#) for details.
offset	a list of offset values if any.
Ntrial	a list of number of trials on Bernoulli experiments. It is useful only for binomialP and binomialPQ variance functions.
power_fixed	a list of logicals indicating if the values of the power parameter should be estimated or not.

weights	Additional weights for the quasi-score function. (Default = NULL).
data	a data frame.
control_initial	
	a list of initial values for the fitting algorithm. If no values are supplied automatic initial values will be provided by the function `mt_initial_values`.
contrasts	extra arguments to passed to `model.matrix`.
control_algorithm	
	a list of arguments to be passed for the fitting algorithm. See `fit_mglm` for details.

Value

mcglm returns an object of class 'mcglm'.

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

Source

Bonat, W. H. and Jorgensen, B. (2016) Multivariate covariance generalized linear models. Journal of Royal Statistical Society - Series C 65:649–675.

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

See Also

`fit_mglm`, `mt_link_function` and `mt_variance_function`.

---

mt_derivative_V_sqrt_beta

Derivatives of V^1/2 with respect to beta.

Description

Compute the derivatives of $V^{1/2}$ matrix with respect to the regression parameters beta.

Usage

```
mt_derivative_V_sqrt_beta(D, D_V_sqrt_mu)
```

Arguments

| D | A matrix. |
| D_V_sqrt_mu | A matrix. |

Value

A list of matrices, containing the derivatives of $V^{1/2}$ with respect to the regression parameters.

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

---

mt_initial_values   *Automatic Initial Values*

Description

This function provides initial values to be used when fitting multivariate generalized linear models by using the function `fit_mglm`. In general the users do not need to use this function, since it is already employed when setting the argument `control_initial = "automatic"` in the `mglm4twin` function. However, if the users want to change some of the initial values, this function can be useful.

Usage

```
mt_initial_values(linear_pred, matrix_pred, link, variance,
                  offset, Ntrial, contrasts, data)
```

Arguments

linear_pred	a list of formula see [formula](#) for details.
matrix_pred	a list of known matrices to be used on the matrix linear predictor. See [mt_matrix_linear_predictor](#) for details.
link	a list of link functions names, see [mglm4twin](#) for details.
variance	a list of variance functions names, see [mglm4twin](#) for details.
offset	a list of offset values if any.
Ntrial	a list of the number of trials on Bernoulli experiments. It is useful only for "binomialP" and "binomialPQ" variance functions.
contrasts	list of contrasts to be used in the [model.matrix](#).
data	data frame.

Details

To obtain initial values for multivariate covariance generalized linear models the function `mt_initial_values` fits a generalized linear model (GLM) using the function `glm` with the specified linear predictor and link function for each response variables considering independent observations. The `family` argument is always specified as `quasi`. The link function depends on the specification of the argument `link`. The variance function is always specified as "mu" the only excession appears when using `variance = "constant"` then the family argument in the `glm` function is specified as `quasi(link = link, variance = "constant")`. The estimated value of the dispersion parameter from the `glm` function is used as initial value for the first component of the matrix

linear predictor, for all other components the value zero is used. The value of the power parameter is always started at 1. In the cases of multivariate models the correlation between response variables is always started at 0.

Value

Return a list of initial values to be used while fitting in the `mglm4twin` function.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

`mt_link_function`	*Link Functions*

Description

The `mt_link_function` is a customized call of the make.link function.

Given the name of a link function, it returns a list with two elements. The first element is the inverse of the link function applied on the linear predictor $\mu = g^{-1}(X\beta)$. The second element is the derivative of μ with respect to the regression parameters β. It will be useful when computing the quasi-score function.

Usage

```
mt_link_function(beta, X, offset, link)

mt_logit(beta, X, offset)

mt_probit(beta, X, offset)

mt_cauchit(beta, X, offset)

mt_cloglog(beta, X, offset)

mt_loglog(beta, X, offset)

mt_identity(beta, X, offset)

mt_log(beta, X, offset)

mt_sqrt(beta, X, offset)

mt_invmu2(beta, X, offset)

mt_inverse(beta, X, offset)
```

Arguments

beta	a numeric vector of regression parameters.
X	a design matrix, see model.matrix for details.
offset	a numeric vector of offset values. It will be sum up on the linear predictor as a covariate with known regression parameter equals one ($\mu = g^{-1}(X\beta + offset)$). If no offset is present in the model, set offset = NULL.
link	a string specifying the name of the link function. Options are: ″logit″, ″probit″, ″cauchit″, ″cloglog″, ″loglog″, ″identity″, ″log″, ″sqrt″, ″1/mu^2″ and inverse. A user defined link function can be used (see Details).

Details

The link function is an important component of the multivariate covariance generalized linear models, since it links the expectation of the response variable with the covariates. Let β be a (p x 1) regression parameter vector and X be an (n x p) design matrix. The expected value of the response variable Y is given by

$$E(Y) = g^{-1}(X\beta),$$

where g is the link function and $\eta = X\beta$ is the linear predictor. Let D be a (n x p) matrix whose entries are given by the derivatives of μ with respect to β. Such a matrix will be required for the fitting algorithm. The function mt_link_function returns a list where the first element is μ (n x 1) vector and the second is the D (n x p) matrix. A user defined function can also be used. It must be a function with arguments beta, X and offset (set to NULL if non needed). The function must return a length 2 named list with mu and D elements as a vector and a matrix of proper dimensions.

Value

A list with two elements: mu and D (see Details).

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

See Also

model.matrix, make.link.

Examples

```
x1 <- seq(-1, 1, l = 5)
X <- model.matrix(~ x1)
mt_link_function(beta = c(1,0.5), X = X,
                 offset = NULL, link = 'log')
mt_link_function(beta = c(1,0.5), X = X,
                 offset = rep(10,5), link = 'identity')
```

`mt_matrix_linear_predictor`

Matrix Linear Predictor

Description

Compute the matrix linear predictor. It is an internal function, however, since the concept of matrix linear predictor was proposed recently. I decided let this function visible to the interested reader gets some feeling about how it works.

Usage

```
mt_matrix_linear_predictor(tau, Z)
```

Arguments

tau   a numeric vector of dispersion parameters.

Z   a list of known matrices.

Details

Given a list with a set of known matrices $(Z_0, ..., Z_D)$ the function
`mt_matrix_linear_predictor` returns $U = \tau_0 Z_0 + ... + \tau_D Z_D$.

Value

A matrix.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

Source

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

Bonat, W. H. and Jorgensen, B. (2016) Multivariate covariance generalized linear models. Journal of Royal Statistical Society - Series C 65:649–675.

Examples

```
require(Matrix)
Z0 <- Diagonal(5, 1)
Z1 <- Matrix(rep(1,5)%*%t(rep(1,5)))
Z <- list(Z0, Z1)
mt_matrix_linear_predictor(tau = c(1,0.8), Z = Z)
```

---

mt_variance_function *Variance Functions*

Description

Compute the variance function and its derivatives with respect to regression, dispersion and power parameters.

Usage

```
mt_variance_function(mu, power, Ntrial, variance,
                          derivative_power, derivative_mu)

mt_tweedie(mu, power, Ntrial, derivative_power, derivative_mu)

mt_binomialP(mu, power, Ntrial,
                      derivative_power, derivative_mu)

mt_binomialPQ(mu, power, Ntrial,
                       derivative_power, derivative_mu)

mt_constant(mu, power, Ntrial, derivative_power, derivative_mu)
```

Arguments

mu	a numeric vector. In general the output from `mt_link_function`.
power	a numeric value (tweedie and binomialP) or a vector (binomialPQ) of the power parameters.
Ntrial	number of trials, useful only when dealing with binomial response variables.
variance	a string specifying the name (constant, tweedie, binomialP or binomialPQ) of the variance function.
derivative_power	
	logical if compute (TRUE) or not (FALSE) the derivatives with respect to the power parameter.
derivative_mu	logical if compute (TRUE) or not (FALSE) the derivative with respect to the mu parameter.

Details

The function `mt_variance_function` computes three features related with the variance function. Depending on the logical arguments, the function returns $V^{1/2}$ and its derivatives with respect to the parameters power and mu, respectively. The output is a named list, completely informative about what the function has been computed. For example, if derivative_power = TRUE and derivative_mu = TRUE. The output will be a list, with three elements: V_sqrt, D_V_sqrt_power and D_V_sqrt_mu.

Value

A list with from one to four elements depending on the arguments.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

Source

Bonat, W. H. and Jorgensen, B. (2016) Multivariate covariance generalized linear models. Journal of Royal Statistical Society - Series C 65:649–675.

See Also

[mt_link_function](#).

Examples

```
x1 <- seq(-1, 1, l = 5)
X <- model.matrix(~x1)
mu <- mt_link_function(beta = c(1, 0.5), X = X, offset = NULL,
                        link = "logit")
mt_variance_function(mu = mu$mu, power = c(2, 1), Ntrial = 1,
                      variance = "binomialPQ",
                      derivative_power = TRUE, derivative_mu = TRUE)
```

---

pAIC   *Pseudo Akaike Information Criterion*

Description

Extract the pseudo Akaike information criterion (pAIC) for objects of mglm4twin class.

Usage

```
pAIC(object, verbose = TRUE)
```

Arguments

object	an object or a list of objects representing a model of mglm4twin class.
verbose	logical. Print or not the pAIC value.

Value

Returns the value of the pseudo Akaike information criterion (pAIC).

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

Source

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

See Also

`gof`, `plogLik`, `pKLIC` and `pBIC`.

---

pBIC   *Pseudo Bayesian Information Criterion*

Description

Extract the pseudo Bayesian information criterion (pBIC) for objects of `mglm4twin` class.

Usage

```
pBIC(object, verbose = TRUE)
```

Arguments

object	an object or a list of objects representing a model of `mglm4twin` class.
verbose	logical. Print or not the pBIC value.

Value

Returns the value of the pseudo Bayesian information criterion (pBIC).

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

Source

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

See Also

`gof`, `plogLik`, `pKLIC` and `pAIC`.

---

pKLIC *Pseudo Kullback-Leibler Information Criterion*

Description

Extract the pseudo Kullback-Leibler information criterion (pKLIC) for objects of `mglm4twin` class.

Usage

```
pKLIC(object, verbose = TRUE)
```

Arguments

object   an object or a list of objects representing a model of `mglm4twin` class.

verbose   logical. Print or not the pKLIC value.

Value

Returns the value of the pseudo Kullback-Leibler information criterion.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

Source

Bonat, W. H. (2018). Multiple Response Variables Regression Models in R: The mcglm Package. Journal of Statistical Software, 84(4):1–30.

See Also

`gof`, `plogLik`, `pAIC`, `pBIC` and `pKLIC`.

---

plogLik *Gaussian Pseudo-loglikelihood*

Description

Extract the Gaussian pseudo-loglikelihood (plogLik) value for objects of `mglm4twin` class.

Usage

```
plogLik(object, verbose = TRUE)
```

Arguments

object	an object or a list of objects representing a model of `mglm4twin` class.
verbose	logical. Print or not the plogLik value.

Value

Returns the value of the Gaussian pseudo-loglikelihood.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

---

  print.mglm4twin   *Print*

Description

The default print method for an object of `mglm4twin` class.

Usage

```
## S3 method for class 'mglm4twin'
print(x, ...)
```

Arguments

x	fitted model objects of class `mglm4twin` as produced by `mglm4twin()`.
...	further arguments passed to or from other methods.

Author(s)

Wagner Hugo Bonat, <wbonat@ufpr.br>

See Also

summary.

---

summary.mglm4twin   *Summarizing*

Description

The default summary method for an object of `mglm4twin` class.

Usage

```
## S3 method for class 'mglm4twin'
summary(object, model, biometric = FALSE, ...)
```

Arguments

object	an object of `mglm4twin` class.
model	String indicating twin model.
biometric	Logical. If TRUE biometric measures such as heritability, common environment, genetic correlation etc are printed. (Default = FALSE).
...	additional arguments affecting the summary produced. Note the there is no extra options for mglm4twin object class.

Value

Print a `mglm4twin` object.

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

See Also

`print`.

---

vcov.mglm4twin   *Variance-Covariance Matrix*

Description

Returns the variance-covariance matrix for an object of `mglm4twin` class.

Usage

```
## S3 method for class 'mglm4twin'
vcov(object, model, ...)
```

Arguments

`object`	an object of `mglm4twin` class.
`model`	String indicating twin model.
`...`	additional arguments affecting the summary produced. Note that there is no extra options for mcglm object class.

Value

A variance-covariance matrix.

Author(s)

Wagner Hugo Bonat, `<wbonat@ufpr.br>`

Index