Package ‘gcmr’

October 13, 2022

Version 1.0.3

Priority optional

Title Gaussian Copula Marginal Regression

Author Guido Masarotto and Cristiano Varin

Maintainer Cristiano Varin <cristiano.varin@unive.it>

Description Likelihood inference in Gaussian copula marginal regression models.

Depends R (>= 4.0.0)

Imports graphics, grDevices, stats, utils, betareg, car, Formula, lmtest, nlme, sandwich, sp

License GPL (>= 2)

NeedsCompilation yes

Repository CRAN

Date/Publication 2022-07-18 07:40:02 UTC

R topics documented:

`gcmr-package` *Gaussian Copula Marginal Regression*

Description

Fits Gaussian copula marginal regression models described in Song (2000) and Masarotto and Varin (2012; 2017).

Details

Gaussian copula models are frequently used to extend univariate regression models to the multivariate case. The principal merit of the approach is that the specification of the regression model is conveniently separated from the dependence structure described in the familiar form of the correlation matrix of a multivariate Gaussian distribution (Song 2000). This form of flexibility has been successfully employed in several complex applications including longitudinal data analysis, spatial statistics, genetics and time series. Some useful references can be found in Masarotto and Varin (2012; 2017) and Song et al. (2013).

This package contains R functions that implement the methodology discussed in Masarotto and Varin (2012) and Guolo and Varin (2014). The main function is `gcmr`, which fits Gaussian copula marginal regression models. Inference is performed through a likelihood approach. Computation of the exact likelihood is possible only for continuous responses, otherwise the likelihood function is approximated by importance sampling. See Masarotto and Varin (2017) for details.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Guolo, A. and Varin, C. (2014). Beta regression for time series analysis of bounded data, with application to Canada Google Flu Trends. *The Annals of Applied Statistics* **8**, 74–88.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

Song, P. X.-K. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics* **27**, 305–320.

Song, P. X.-K., Li, M. and Zhang, P. (2013). Copulae in Mathematical and Quantitative Finance. In *Vector Generalized Linear Models: A Gaussian Copula Approach*, 251–276. Springer Berlin Heidelberg.

arma.cormat *ARMA(p,q) Correlation*

Description

Sets ARMA(p,q) correlation in Gaussian copula regression models.

Usage

```
arma.cormat(p, q)
```

Arguments

p	order of the autoregressive component.
q	order of the moving average component.

Value

An object of class `cormat.gcmr` representing a correlation matrix with ARMA(p,q) structure.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`.

cluster.cormat *Longitudinal/Clustered Data Correlation*

Description

Sets longitudinal/clustered data correlation in Gaussian copula regression models.

Usage

```
cluster.cormat(id, type = c("independence", "ar1", "ma1",
        "exchangeable", "unstructured"))
```

Arguments

`id`	subject id. This is a vector of the same lenght of the number of observations. Please note that data must be sorted in way that observations from the same cluster are contiguous.
`type`	a character string specifying the correlation structure. At the moment, the following are implemented:

`independence`	working independence.
`ar1`	autoregressive of order 1.
`ma1`	moving average of order 1.
`exchangeable`	exchangeable.
`unstructured`	unstructured.

Details

The correlation matrices are inherited from the `nlme` package (Pinheiro and Bates, 2000).

Value

An object of class `cormat.gcmr` representing a correlation matrix for longitudinal or clustered data.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer.

See Also

`gcmr`, `nlme`.

`cormat.gcmr` *Correlation Matrices for Gaussian Copula Regression Models*

Description

Class of correlation matrices available in the `gcmr` package.

Value

At the moment, the following are implemented:

`ind.cormat`	working independence.
`arma.cormat`	ARMA(p,q).
`cluster.cormat`	longitudinal/clustered data.
`matern.cormat`	Matern spatial correlation.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`, `ind.cormat`, `arma.cormat`, `cluster.cormat`, `matern.cormat`.

`epilepsy` *Epilitic Seizures Data*

Description

Longitudinal study on epilitic seizures (Thall and Vail, 1990; Diggle et al. 2002). The data consist into 59 individuals with five observations each: The baseline eight-week interval and measurements collected at subsequent visits every two-week.

Usage

```
data(epilepsy)
```

Format

`id`	patient's id .
`age`	patient's age.
`trt`	indicator if the patient is treated with progabide (`1`) or with placebo (`2`).
`counts`	number of epileptic seizures.
`time`	observation period in weeks (`8` for baseline and `2` for subsequent visits).
`visit`	indicator if observation at baseline (`0`) or subsequent visit (`1`).

Source

Thall, P.F. and Vail S.C. (1990). Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46**, 657–671.

References

Diggle, P.J., Heagerty, P., Liang, K.Y. and Zeger, S.L. (2002). *Analysis of Longitudinal Data*. Oxford: Oxford University Press. Second edition.

`gaussian.marg` *Marginals in Gaussian Copula Marginal Regression Models*

Description

These functions set the marginals in Gaussian copula marginal regression models.

Usage

```
beta.marg(link = "logit")
binomial.marg(link = "logit")
Gamma.marg(link = "inverse")
gaussian.marg(link = "identity")
negbin.marg(link = "log")
poisson.marg(link = "log")
weibull.marg(link = "log")
```

Arguments

`link`	a specification for the model link function. See `family` for the special case of generalized linear models.

Details

Beta marginals specified by `beta.marg` are parametrized in terms of mean and dispersion as in `betareg`. See Cribari-Neto and Zeileis (2010) and Ferrari and Cribari-Neto (2004).

For binomial marginals specified by `binomial.marg`, the response is specified as a factor when the first level denotes failure and all others success or as a two-column matrix with the columns giving the numbers of successes and failures.

Negative binomial marginals implemented in `negbin.marg` are parametrized such that $var(Y) = E(Y) + kE(Y)^2$.

For back-compatibility with previous versions of the `gcmr` package, short names for the marginals `bn.marg`, `gs.marg`, `nb.marg`, and `ps.marg` remain valid as an alternative to (preferred) longer versions `binomial.marg`, `gaussian.marg`, `negbin.marg`, and `poisson.marg`.

Value

An object of class `marginal.gcmr` representing the marginal component.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software* **34**, 1–24.

Ferrari, S.L.P. and Cribari-Neto, F. (2004). Beta regression for modeling rates and proportions. *Journal of Applied Statistics* **31** (7), 799–815.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`, `betareg`.

---

gcmr — *Fitting Gaussian Copula Marginal Regression Models by Maximum (Simulated) Likelihood.*

Description

Fits Gaussian copula marginal regression models by maximum (simulated) likelihood.

Usage

```
gcmr(formula, data, subset, offset, marginal,
     cormat, start, fixed, options=gcmr.options(...), model=TRUE,...)

gcmr.fit(x=rep(1,NROW(y)), y, z=NULL, offset=NULL,
 marginal, cormat, start, fixed, options=gcmr.options())
```

Arguments

`formula`	a symbolic description of the model to be fitted of type `y ~ x` or `y ~ x	z`, for details see below.
`data`	an optional data frame, list or environment (or object coercible by `as.data.frame` to a data frame) containing the variables in the model. If not found in `data`, the variables are taken from `environment(formula)`.
`subset`	an optional vector specifying a subset of observations to be used in the fitting process.
`offset`	optional numeric vector with an a priori known component to be included in the linear predictor for the mean. When appropriate, offset may also be a list of two offsets for the mean and precision equation, respectively.

x	design matrix.
y	vector of observations.
z	optional design matrix for the dispersion/shape.
marginal	an object of class marginal.gcmr specifying the marginal part of the model.
cormat	an object of class cormat.gcmr representing the correlation matrix of the errors.
start	optional numeric vector with starting values for the model parameters.
fixed	optional numeric vector of the same length as the total number of parameters. If supplied, only NA entries in fixed will be varied.
options	list of options passed to function gcmr.options.
model	logical. If TRUE, then the model frame is returned.
...	arguments passed to gcmr.options.

Details

Function gcmr computes maximum likelihood estimation in Gaussian copula marginal regression models. Computation of the exact likelihood is possible only for continuous responses, otherwise the likelihood function is approximated by importance sampling. See Masarotto and Varin (2012; 2017) for details.

Standard formula `y ~ x1 + x2` indicates that the mean response is modelled as a function of covariates `x1` and `x2` through an appropriate link function. Extended formula `y ~ x1 + x2 | z1 + z2` indicates that the dispersion (or the shape) parameter of the marginal distribution is modelled as a function of covariates `z1` and `z2`. Dispersion (or shape) parameters are always modelled on logarithm scale. The model specification is inspired by beta regression as implemented in betareg (Cribari-Neto and Zeileis, 2010) through extended Formula objects (Zeileis and Croissant, 2010).

For binomial marginals specified by binomial.marg the response is specified as a factor when the first level denotes failure and all others success or as a two-column matrix with the columns giving the numbers of successes and failures.

`gcmr.fit` is the workhorse function: it is not normally called directly but can be more efficient where the response vector and design matrix have already been calculated.

Value

An object of class "gcmr" with the following components:

estimate	the maximum likelihood estimate.
maximum	the maximum likelihood value.
hessian	(minus) the Hessian at the maximum likelihood estimate.
jac	the Jacobian at the maximum likelihood estimate.
fitted.values	the fitted values.
marginal	the marginal model used.
cormat	the correlation matrix used.
fixed	the numeric vector indicating which parameters are constants.
ibeta	the indices of marginal parameters.

igamma	the indices of dependence parameters.
nbeta	the number of marginal parameters.
ngamma	the number of dependence parameters.
options	the fitting options used, see gcmr.options.
call	the matched call.
formula	the model formula.
terms	the terms objects for the fitted model.
levels	the levels of the categorical regressors.
model	the model frame, returned only if model=TRUE.
contrasts	the contrasts corresponding to levels.
y	the y vector used.
x	the model matrix used for the mean response.
z	the (optional) model matrix used for the dispersion/shape.
offset	the offset used.
n	the number of observations.
not.na	the vector of binary indicators of the available observations (not missing).

Functions coefficients, logLik, fitted, vcov.gcmr and residuals.gcmr can be used to extract various useful features of the value returned by gcmr. Function plot.gcmr produces various diagnostic plots for fitted gcmr objects.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Cribari-Neto, F. and Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software* **34**, 1–24.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

Rocha, A.V. and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *Test* **18**, 529–545.

Zeileis, A. and Croissant, Y. (2010). Extended model formulas in R: Multiple parts and multiple responses. *Journal of Statistical Software* **34**, 1–13.

See Also

cormat.gcmr, marginal.gcmr, gcmr.options, Formula, betareg.

Examples

```
## negative binomial model for longitudinal data
data(epilepsy)
gcmr(counts ~ offset(log(time)) + visit + trt + visit:trt, data = epilepsy,
subset = (id != 49), marginal = negbin.marg, cormat = cluster.cormat(id, "ar1"),
options=gcmr.options(seed=123, nrep=100 ))
## Hidden Unemployment Rate (HUR) data (Rocha and Cribari-Neto, 2009)
## beta regression with ARMA(1,3) errors
data(HUR)
trend <- scale(time(HUR))
gcmr(HUR ~ trend | trend, marginal = beta.marg, cormat = arma.cormat(1, 3))
```

---

gcmr.options *Setting Options for Fitting Gaussian Copula Marginal Regression Models*

Description

Sets options that affect the fitting of Gaussian copula marginal regression models.

Usage

```
gcmr.options(seed = round(runif(1, 1, 1e+05)), nrep = c(100, 1000),
    no.se = FALSE, method = c("BFGS", "Nelder-Mead", "CG"), ...)
```

Arguments

`seed`	seed of the pseudorandom generator used in the importance sampling algorithm for likelihood approximation in case of discrete responses.
`nrep`	Monte Carlo size of the importance sampling algorithm for likelihood approximation in case of discrete responses. `nrep` can be a vector so that the model is fitted with a sequence of different Monte Carlo sizes. In this case, the starting values for optimization of the likelihood are taken from the previous fitting. A reasonable strategy is to fit the model with a small Monte Carlo size to obtain sensible starting values and then refit with a larger Monte Carlo size. The default value is `100` for the first optimization and `1000` for the second and definitive optimization.
`no.se`	logical. Should standard errors be computed and returned or not?
`method`	a character string specifying the method argument passed to `optim`. The default optimization routine is the quasi-Newton algorithm `BFGS`. See `optim` for details.
`...`	arguments passed to `optim`.

Value

A list containing the options.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`

HUR *Hidden Unemployment in Sao Paulo*

Description

Rate of hidden unemployment due to substandard work conditions in Sao Paulo, Brazil (Rocha and Cribari-Neto, 2009).

Usage

```
data(HUR)
```

Source

Institute of Applied Economic Research (Ipea), Brazil. Data obtained from the IPEAdata website `http://www.ipeadata.gov.br`.

References

Rocha, A.V. and Cribari-Neto, F. (2009). Beta autoregressive moving average models. *Test* **18**, 529–545.

`ind.cormat` *Working Independence Correlation*

Description

Sets working independence correlation in Gaussian copula marginal regression models.

Usage

```
ind.cormat()
```

Value

An object of class `cormat.gcmr` representing an identity correlation matrix.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`.

`malaria` *Gambia Malaria Data*

Description

Malaria prevalence in children in Gambia. The data are constructed from the `gambia` dataframe in the `geoR` package (Diggle and Ribeiro, 2007) by village aggregation.

Usage

```
data(malaria)
```

Format

A data frame with the 65 observations with the following variables

`x`	x-coordinate of the village (UTM).
`y`	y-coordinate of the village (UTM).
`cases`	number of sampled children with malaria in each village.
`size`	number of sampled children in each village.
`age`	mean age of the sampled children in each village.
`netuse`	frequency of sampled children who regularly sleep under a bed-net in each village.
`treated`	frequency of sampled children whose bed-net is treated.
`green`	measure of vegetation green-ness in the immediate vicinity of the village.
`phc`	indicator variable denoting the presence (1) or absence (0) of a health center in the village.
`area`	indicator of the village area (Diggle et al., 2002).

Source

Diggle, P.J. and Ribeiro Jr, P.J. (2007). *Model Based Geostatistics*. New York: Springer.

References

Thomson, M., Connor, S., D Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M. and Greenwood, B. (1999). Predicting malaria infection in Gambian children from satellite data and bednet use surveys: the importance of spatial correlation in the interpretation of results. *American Journal of Tropical Medicine and Hygiene* **61**, 2–8.

Diggle, P., Moyeed, R., Rowlingson, B. and Thomson, M. (2002). Childhood malaria in The Gambia: a case-study in model-based geostatistics, *Applied Statistics* **51**, 493–506.

Examples

```
data(malaria)
```

`marginal.gcmr` *Marginals for Gaussian Copula Marginal Regression*

Description

Class of marginals available in the `gcmr` library.

Value

At the moment, the following are implemented:

`beta.marg`	beta marginals.
`binomial.marg`	binomial marginals.
`Gamma.marg`	Gamma marginals.
`gaussian.marg`	Gaussian marginals.

negbin.marg	negative binomial marginals.
poisson.marg	Poisson marginals.
weibull.marg	Weibull marginals.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

gcmr, beta.marg, binomial.marg, gaussian.marg, Gamma.marg, negbin.marg, poisson.marg, weibull.marg.

---

matern.cormat *Matern Spatial Correlation*

Description

Sets a Matern spatial correlation matrix in Gaussian copula marginal regression models.

Usage

```
matern.cormat(D, alpha = 0.5)
```

Arguments

D	matrix with values of the distances between pairs of data locations.
alpha	value of the shape parameter of the Matern correlation class. The default alpha = 0.5 corresponds to an exponential correlation model.

Details

The Mat\'ern correlation function is inherited from the geoR package (Diggle and Ribeiro, 2007).

Value

An object of class cormat.gcmr representing a Matern correlation matrix.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Diggle, P. and Ribeiro, P.J. (2007). *Model-based Geostatistics*. Springer.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

gcmr.

---

plot.gcmr *Plot Diagnostics for Gaussian Copula Marginal Regression*

Description

Various types of diagnostic plots for Gaussian copula regression.

Usage

```
## S3 method for class 'gcmr'
plot(x, which = if (!time.series) 1:4 else c(1, 3, 5, 6),
    caption = c("Residuals vs indices of obs.", "Residuals vs linear predictor",
        "Normal plot of residuals", "Predicted vs observed values",
        "Autocorrelation plot of residuals", "Partial ACF plot of residuals"),
     main = "", ask = prod(par("mfcol")) < length(which) && dev.interactive(),
     level = 0.95, col.lines = "gray",
time.series = inherits(x$cormat, "arma.gcmr"), ...)
```

Arguments

`x`	a fitted model object of class `gcmr`.
`which`	select one, or more, of the six available plots. The default choice adapts to the correlation structure and selects four plots depending on the fact that the data are a regular time series or not.
`caption`	captions to appear above the plots.
`main`	title to each plot in addition to the above caption.
`ask`	if `TRUE`, then the user is asked before each plot.
`level`	confidence level in the normal probability plot. The default is `0.95`.
`col.lines`	color for lines. The default is `"gray"`.

`time.series` if `TRUE`, four plots suitable for time series data are displayed. The default is `TRUE` when the correlation matrix corresponds to that of ARMA(p,q) process and `FALSE` otherwise.

`...` other parameters to be passed through to plotting functions.

Details

The plot method for `gcmr` objects produces six types of diagnostic plots selectable through the `which` argument. Available choices are: Quantile residuals vs indices of the observations (`which=1`); Quantile residuals vs linear predictor (`which=2`); Normal probability plot of quantile residuals (`which=3`); Fitted vs observed values (`which=4`); Autocorrelation plot of quantile residuals (`which=5`); Partial autocorrelation plot of quantile residuals (`which=6`). The latter two plots make sense for regular time series data only.

The normal probability plot is computed via function `qqPlot` from the package `car` (Fox and Weisberg, 2011).

Author(s)

Guido Masarotto and Cristiano Varin.

References

Fox, J. and Weisberg, S. (2011). *An R Companion to Applied Regression*. Second Edition. Thousand Oaks CA: Sage.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`.

Examples

```
## beta regression with ARMA(1,3) errors
data(HUR)
trend <- scale(time(HUR))
m <- gcmr(HUR ~ trend | trend, marginal = beta.marg, cormat = arma.cormat(1, 3))
## normal probability plot
plot(m,  3)
## autocorrelation function of residuals
plot(m,  5)
```

polio *Polio Time Series*

Description

Time series of Polio incidences in U.S.A. from 1970 to 1983.

Usage

```
data(polio)
```

Format

A data frame with the 168 monthly observations (from January 1970 to December 1983) with the following variables

`y`	time series of polio incidences.
`t*10^( -3 )`	linear trend multiplied by factor $10^{(-3)}$.
`cos( 2*pi*t/12 )`	cosine annual seasonal component.
`sin( 2*pi*t/12 )`	sine annual seasonal component.
`cos( 2*pi*t/6 )`	cosine semi-annual seasonal component.
`sin( 2*pi*t/6 )`	sine semi-annual seasonal component.

Source

Zeger, S.L. (1988). A regression model for time series of counts. *Biometrika* **75**, 822–835.

Examples

```
data(polio)
```

profile.gcmr *Profile Log-Likelihood for Gaussian Copula Marginal Regression Models*

Description

Computes the profile log-likelihood for mean response parameters of a Gaussian copula marginal regression model.

Usage

```
## S3 method for class 'gcmr'
profile(fitted, which, low, up, npoints = 10,
        display = TRUE, alpha = 0.05, progress.bar = TRUE, ...)
```

Arguments

fitted	a fitted Gaussian copula marginal regression model of class `gcmr`.
which	the index of the regression parameter which should be profiled.
low	the lower limit used in computation of the profile log-likelihood. If this is `missing`, then the lower limit is set equal to the estimate minus three times its standard error.
up	the upper limit used in computation of the profile log-likelihood. If this is `missing`, then the upper limit is set equal to the estimate plus three times its standard error.
npoints	number of points used in computation of the profile log-likelihood. Default is `10`.
display	should the profile log-likelihood be displayed or not? default is `TRUE`.
alpha	the significance level, default is `0.05`.
progress.bar	logical. If TRUE, a text progress bar is displayed.
...	further arguments passed to `plot`.

Details

If the display is requested, then the profile log-likelihood is smoothed by cubic spline interpolation.

Value

A list with the following components:

points	points at which the profile log-likelihood is evaluated.
profile	values of the profile log-likelihood.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

See Also

`gcmr`

Examples

```
## spatial binomial data
## Not run:
data(malaria)
D <- sp::spDists(cbind(malaria$x, malaria$y))/1000
m <- gcmr(cbind(cases, size-cases) ~ netuse+I(green/100)+phc, data=malaria,
marginal=binomial.marg, cormat=matern.cormat(D), options=gcmr.options(seed=987))
prof <- profile(m, which = 2)
prof

## End(Not run)
```

`residuals.gcmr` *Quantile Residuals for Gaussian Copula Marginal Regression*

Description

Computes various type of quantile residuals for validation of a fitted Gaussian copula marginal regression model, as described in Masarotto and Varin (2012; 2017).

Usage

```
## S3 method for class 'gcmr'
residuals(object, type=c("conditional","marginal"),
          method=c("random","mid"),...)
```

Arguments

`object`	an object of class `gcmr`, typically the result of a call to `gcmr`.
`type`	the type of quantile residuals which should be returned. The alternatives are: `"conditional"` (default) and `"marginal"`.
`method`	different methods available for quantile residuals in case of discrete responses: `"random"` for randomized quantile residuals (default), and `"mid"` for mid interval quantile residuals as defined in Zucchini and MacDonald (2009).
`...`	further arguments passed to or from other methods.

Details

Quantile residuals are defined in Dunn and Smyth (1996). Two different types are available:

`conditional`	quantile residuals that account for the dependence.
`marginal`	quantile residuals that do not account for the dependence.

Conditional quantile residuals are normal quantiles of Rosenblatt (1952) transformations and they are appropriate for validation of the marginal regression models discussed in Masarotto and Varin (2012; 2017). If the responses are discrete, then the conditional quantile residuals are not well defined. This difficulty is overcame by randomized quantile residuals available through option `method="random"`. Alternatively, Zucchini and MacDonald (2009) suggest the use of mid interval quantile residuals (`method="mid"`).

Note

Differently from randomized quantile residuals, mid quantile residuals are **not** realizations of incorrelated standard normal variables under model conditions.

It is appropriate to inspect several sets of randomized quantile residuals before to take a decision about the model.

See Masarotto and Varin (2012; 2017) for more details.

Author(s)

Guido Masarotto and Cristiano Varin.

References

Dunn, P.K. and Smyth, G.K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**, 236–244.

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* **6**, 1517–1549.

Masarotto, G. and Varin C. (2017). Gaussian Copula Regression in R. *Journal of Statistical Software*, **77**(8), 1–26.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics* **23**, 470–472.

Zucchini, W. and MacDonald, I.L. (2009). *Hidden Markov Models for Time Series*. Chapman and Hall/CRC.

See Also

`gcmr`

Examples

```
## spatial binomial data
## Not run:
data(malaria)
D <- sp::spDists(cbind(malaria$x, malaria$y))/1000
m <- gcmr(cbind(cases, size-cases) ~ netuse+I(green/100)+phc, data=malaria,
marginal=binomial.marg, cormat=matern.cormat(D))
res <- residuals(m)
## normal probability plot
qqnorm(res)
qqline(res)
## or better via plot.gcmr
plot(m, which = 3)

## End(Not run)
```

`scotland` *Scotland Lip Cancer Data*

Description

Male lip cancer in Scotland counties between 1975-1980.

Usage

```
data(scotland)
```

Format

A data frame with the 56 observations with the following variables

`observed`	observed cases in each county.
`expected`	expected cases in each county.
`AFF`	proportion of the population employed in agriculture, fishing, or forestry.
`latitude`	county latitude.
`longitude`	county longitude.

Source

Waller, L.A. and Gotway, C.A. (2004). *Applied Spatial Statistics for Public Health Data*. New York: John Wiley and Sons.

References

Clayton D. and Kaldor J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* **43**, 671–681.

Examples

```
data(scotland)
```

`summary.gcmr` *Methods for gcmr Objects*

Description

Methods for extracting information from fitted beta regression model objects of class `"gcmr"`.

Usage

```
## S3 method for class 'gcmr'
summary(object, ...)

## S3 method for class 'gcmr'
coef(object, ...)
## S3 method for class 'gcmr'
vcov(object, ...)
## S3 method for class 'gcmr'
bread(x, ...)
## S3 method for class 'gcmr'
estfun(x, ...)
```

Arguments

`object, x`	a fitted marginal regression model of class `gcmr`.
`...`	additional arguments, but currently not used.

Value

The function `summary.gcmr` returns an object of class "summary.glm", a list with some components of the `gcmr` object, plus

`coefficients`	a list with components `marginal` and `copula` containing the maximum likelihood estimates of the marginal and Gaussian copula parameters, respectively.
`aic`	Akaike Information Criterion.

Function `coef` returns the estimated coefficients and `vcov` their variance-covariance matrix. Functions `bread` and `estfun` extract the components of the robust sandwich variance matrix that can be computed with the `sandwich` package (Zeileis, 2004; 2006).

Author(s)

Guido Masarotto and Cristiano Varin.

References

Zeileis, A. (2004). Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software* **11**, issue 10.

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software* **16**, issue 9.

See Also

`bread`, `estfun`, `gcmr`, `sandwich`.

Examples

```
data(epilepsy)
fit <- gcmr(counts ~ offset(log(time)) + visit + trt + visit:trt, data = epilepsy,
subset = (id != 49), marginal = negbin.marg, cormat = cluster.cormat(id, "ar1"),
options=gcmr.options(seed=123, nrep=c(25,100) ))
summary(fit)
```

Index