

Package ‘cquad’

October 10, 2020

Type Package

Title Conditional Maximum Likelihood for Quadratic Exponential Models
for Binary Panel Data

Version 2.2

Date 2020-10-10

Author Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona ``Politecnica delle Marche''), Francesco Valentini (University of Ancona ``Politecnica delle Marche'')

Maintainer Francesco Bartolucci <francesco.bartolucci@unipg.it>

Description Estimation, based on conditional maximum likelihood, of the quadratic exponential model proposed by Bartolucci, F. & Nigro, V. (2010, *Econometrica*) <DOI:10.3982/ECTA7531> and of a simplified and a modified version of this model. The quadratic exponential model is suitable for the analysis of binary longitudinal data when state dependence (further to the effect of the covariates and a time-fixed individual intercept) has to be taken into account. Therefore, this is an alternative to the dynamic logit model having the advantage of easily allowing conditional inference in order to eliminate the individual intercepts and then getting consistent estimates of the parameters of main interest (for the covariates and the lagged response). The simplified version of this model does not distinguish, as the original model does, between the last time occasion and the previous occasions. The modified version formulates in a different way the interaction terms and it may be used to test in a easy way state dependence as shown in Bartolucci, F., Nigro, V. & Pigni, C. (2018, *Econometric Reviews*) <DOI:10.1080/07474938.2015.1060039>. The package also includes estimation of the dynamic logit model by a pseudo conditional estimator based on the quadratic exponential model, as proposed by Bartolucci, F. & Nigro, V. (2012, *Journal of Econometrics*) <DOI:10.1016/j.jeconom.2012.03.004>. For large time dimensions of the panel, the computation of the proposed models involves a recursive function adapted from Krailo M. D., & Pike M. C. (1984, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*).

License GPL (>= 2)

Depends R (>= 2.0.0), MASS, plm

NeedsCompilation no

Repository CRAN

Date/Publication 2020-10-10 14:10:03 UTC

R topics documented:

| | |
|------------------|-----------|
| cquad-package | 2 |
| cquad | 3 |
| cquad_basic | 5 |
| cquad_equ | 6 |
| cquad_ext | 7 |
| cquad_pseudo | 9 |
| data_sim | 10 |
| print.cquad | 11 |
| quasi_sym | 11 |
| quasi_sym_equ | 12 |
| quasi_sym_pseudo | 13 |
| sim_panel_logit | 14 |
| sq | 15 |
| summary.cquad | 16 |
| Index | 17 |

| | |
|---------------|--|
| cquad-package | <i>Conditional Maximum Likelihood for Quadratic Exponential Models for Binary Panel Data</i> |
|---------------|--|

Description

Estimation, based on conditional maximum likelihood, of the quadratic exponential model proposed by Bartolucci & Nigro (2010) and of a simplified and a modified version of this model. The quadratic exponential model is suitable for the analysis of binary longitudinal data when state dependence (further to the effect of the covariates and a time-fixed individual intercept) has to be taken into account. Therefore, this is an alternative to the dynamic logit model having the advantage of easily allowing conditional inference in order to eliminate the individual intercepts and then getting consistent estimates of the parameters of main interest (for the covariates and the lagged response). The simplified version of this model does not distinguish, as the original model does, between the last time occasion and the previous occasions. The modified version formulates in a different way the interaction terms and it may be used to test in a easy way state dependence as shown in Bartolucci, Nigro & Pignini (2018). The package also includes estimation of the dynamic logit model by a pseudo conditional estimator based on the quadratic exponential model, as proposed by Bartolucci & Nigro (2012).

Details

| | |
|----------|------------|
| Package: | cquad |
| Type: | Package |
| Version: | 2.0 |
| Date: | 2019-03-26 |
| License: | GPL (>= 2) |

Author(s)

Francesco Bartolucci (University of Perugia, IT), Claudia Pigni (University of Perugia, IT), Francesco Valentini (University of Ancona "Politecnica delle Marche")

Maintainer: Francesco Bartolucci <francesco.bartolucci@unipg.it>

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.

Bartolucci, F. and Nigro, V. (2012). Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data, *Journal of Econometrics*, **170**, 102-116.

Bartolucci, F. and Pigni, C. (2017). cquad: An R and Stata package for conditional maximum likelihood estimation of dynamic binary panel data models, *Journal of Statistical Software*, **78**, 1-26, doi:10.18637/jss.v078.i07.

Bartolucci, F., Nigro, V., & Pigni, C. (2018). Testing for state dependence in binary panel data with individual covariates by a modified quadratic exponential model. *Econometric Reviews*, **37(1)**, 61-88.

Cox, D. R. (1972), The Analysis of multivariate binary data, *Applied Statistics*, **21**, 113-120.

Examples

```
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
# static model
out1 = cquad(y~X1+X2,data_sim)
# dynamic model
out2 = cquad(y~X1+X2,data_sim,dyn=TRUE)
```

cquad

Interface for functions fitting different versions of cquad

Description

Fit by conditional maximum likelihood each of the models in cquad package.

Usage

```
cquad(formula, data, index = NULL, model = c("basic", "equal", "extended", "pseudo"),
      w = rep(1, n), dyn = FALSE, Ttol=10)
```

Arguments

| | |
|---------|---|
| formula | formula with the same syntax as in plm package |
| data | data.frame or pdata.frame |
| index | to denote panel structure as in plm package |
| model | type of model = "basic", "equal", "extended", "pseudo" |
| w | vector of weights (optional) |
| dyn | TRUE if in the dynamic version; FALSE for the static version (by default) |
| Ttol | Threshold individual observations that activates the recursive algorithm (default=10) |

Value

| | |
|--------------|---|
| formula | formula defining the model |
| lk | conditional log-likelihood value |
| coefficients | estimate of the regression parameters |
| vcov | asymptotic variance-covariance matrix for the parameter estimates |
| scv | matrix of individual scores |
| J | Hessian of the log-likelihood function |
| se | standard errors |
| ser | robust standard errors |
| Tv | number of time occasions for each unit |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

Examples

```
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
# basic (static) model
out1 = cquad(y~X1+X2,data_sim)
summary(out1)
# basic (dynamic) model
out2 = cquad(y~X1+X2,data_sim,dyn=TRUE)
summary(out2)
# equal model
out3 = cquad(y~X1+X2,data_sim,model="equal")
summary(out3)
# extended model
out4 = cquad(y~X1+X2,data_sim,model="extended")
summary(out4)
# psuedo CML for dynamic model
out5 = cquad(y~X1+X2,data_sim,model="pseudo")
summary(out5)
```

| | |
|-------------|---|
| cquad_basic | <i>Conditional maximum likelihood estimation of the basic quadratic exponential model</i> |
|-------------|---|

Description

Fit by conditional maximum likelihood a simplified version of the model for binary longitudinal data proposed by Bartolucci & Nigro (2010); see also Cox (1972).

Usage

```
cquad_basic(id, yv, X = NULL, be = NULL, w = rep(1, n), dyn =
FALSE, Ttol=10)
```

Arguments

| | |
|------|---|
| id | list of the reference unit of each observation |
| yv | corresponding vector of response variables |
| X | corresponding matrix of covariates (optional) |
| be | initial vector of parameters (optional) |
| w | vector of weights (optional) |
| dyn | TRUE if in the dynamic version; FALSE for the static version (by default) |
| Ttol | Threshold individual observations that activates the recursive algorithm (default=10) |

Value

| | |
|--------------|--|
| formula | formula defining the model |
| lk | conditional log-likelihood value |
| coefficients | estimate of the regression parameters (including for the lag-response) |
| vcov | asymptotic variance-covariance matrix for the parameter estimates |
| scv | matrix of individual scores |
| J | Hessian of the log-likelihood function |
| se | standard errors |
| ser | robust standard errors |
| Tv | number of time occasions for each unit |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, pp. 719-733.

Cox, D. R. (1972), The Analysis of multivariate binary data, *Applied Statistics*, **21**, 113-120.

Examples

```
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
id = data_sim$id; yv = data_sim$y; X = cbind(X1=data_sim$X1,X2=data_sim$X2)
# static model
out1 = cquad_basic(id,yv,X,Ttol=10)
summary(out1)
# dynamic model
out2 = cquad_basic(id,yv,X,dyn=TRUE,Ttol=10)
summary(out2)
```

cquad_equ

Conditional maximum likelihood estimation for the modified version of the quadratic exponential model (to test for state dependence)

Description

Fit by conditional maximum likelihood a modified version of the model for binary longitudinal data proposed by Bartolucci & Nigro (2010), in which the interaction terms have an extended form. This modified version is used to test for state dependence as described in Bartolucci et al. (2018).

Usage

```
cquad_equ(id, yv, X = NULL, be = NULL, w = rep(1, n), Ttol=10)
```

Arguments

| | |
|------|---|
| id | list of the reference unit of each observation |
| yv | corresponding vector of response variables |
| X | corresponding matrix of covariates (optional) |
| be | initial vector of parameters (optional) |
| w | vector of weights (optional) |
| Ttol | Threshold individual observations that activates the recursive algorithm (default=10) |

Value

| | |
|--------------|--|
| formula | formula defining the model |
| lk | conditional log-likelihood value |
| coefficients | estimate of the regression parameters (including for the lag-response) |
| vcov | asymptotic variance-covariance matrix for the parameter estimates |
| scv | matrix of individual scores |
| J | Hessian of the log-likelihood function |
| se | standard errors |
| ser | robust standard errors |
| Tv | number of time occasions for each unit |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Perugia), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.

Bartolucci, F., Nigro, V., & Pigni, C. (2018). Testing for state dependence in binary panel data with individual covariates by a modified quadratic exponential model. *Econometric Reviews*, **37**(1), 61-88.

Examples

```
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
id = data_sim$id; yv = data_sim$y; X = cbind(X1=data_sim$X1,X2=data_sim$X2)

out = cquad_equ(id,yv,X,Ttol=10)
```

| | |
|-----------|--|
| cquad_ext | <i>Conditional maximum likelihood estimation of the quadratic exponential model for panel data</i> |
|-----------|--|

Description

Fit by conditional maximum likelihood the model for binary longitudinal data proposed by Bartolucci & Nigro (2010).

Usage

```
cquad_ext(id, yv, X = NULL, be = NULL, w = rep(1, n),Ttol=10)
```

Arguments

| | |
|------|---|
| id | list of the reference unit of each observation |
| yv | corresponding vector of response variables |
| X | corresponding matrix of covariates (optional) |
| be | initial vector of parameters (optional) |
| w | vector of weights (optional) |
| Ttol | Threshold individual observations that activates the recursive algorithm (default=10) |

Value

| | |
|--------------|--|
| formula | formula defining the model |
| lk | conditional log-likelihood value |
| coefficients | estimate of the regression parameters (including for the lag-response) |
| vcov | asymptotic variance-covariance matrix for the parameter estimates |
| scv | matrix of individual scores |
| J | Hessian of the log-likelihood function |
| se | standard errors |
| ser | robust standard errors |
| Tv | number of time occasions for each unit |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator. *Econometrica*, **78**, pp. 719-733.

Examples

```
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
id = data_sim$id; yv = data_sim$y; X = cbind(X1=data_sim$X1,X2=data_sim$X2)
# static model
out = cquad_ext(id,yv,X,Ttol=10)
summary(out)
```

| | |
|--------------|--|
| cquad_pseudo | <i>Pseudo conditional maximum likelihood estimation of the dynamic logit model</i> |
|--------------|--|

Description

Estimate the dynamic logit model for binary longitudinal data by the pseudo conditional maximum likelihood method proposed by Bartolucci & Nigro (2012).

Usage

```
cquad_pseudo(id, yv, X = NULL, be = NULL, w = rep(1,n), Ttol=10)
```

Arguments

| | |
|------|---|
| id | list of the reference unit of each observation |
| yv | corresponding vector of response variables |
| X | corresponding matrix of covariates (optional) |
| be | initial vector of parameters (optional) |
| w | vector of weights (optional) |
| Ttol | Threshold individual observations that activates the recursive algorithm (default=10) |

Value

| | |
|--------------|--|
| formula | formula defining the model |
| lk | conditional log-likelihood value |
| coefficients | estimate of the regression parameters (including for the lag-response) |
| vcov | asymptotic variance-covariance matrix for the parameter estimates |
| scv | matrix of individual scores |
| J | Hessian of the log-likelihood function |
| se | standard errors |
| se2 | robust standard errors that also take into account the first step |
| Tv | number of time occasions for each unit |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.

Bartolucci, F. and Nigro, V. (2012), Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data, *Journal of Econometrics*, **170**, 102-116.

Examples

```
## Not run:
# example based on simulated data
data(data_sim)
data_sim = data_sim[1:500,] # to speed up the example, remove otherwise
id = data_sim$id; yv = data_sim$y; X = cbind(X1=data_sim$X1,X2=data_sim$X2)
# estimate dynamic logit model
out = cquad_pseudo(id,yv,X, Ttol=10)
summary(out)

## End(Not run)
```

| | |
|----------|--------------------------|
| data_sim | <i>Simulated dataset</i> |
|----------|--------------------------|

Description

It contains a dataset simulated from the dynamic logit model

Usage

```
data(data_sim)
```

Format

The observations are for 1000 sample units at 5 five time occasions:

id list of the reference unit of each observation

time number of the time occasion

X1 first covariate

X2 second covariate

y response

Examples

```
data(data_sim)
head(data_sim)
```

| | |
|-------------|-------------------------------------|
| print.cquad | <i>Print output for class cquad</i> |
|-------------|-------------------------------------|

Description

Print output for class cquad and output provided by cquad_basic, cquad_equ, cquad_ext, cquad_pseudo

Usage

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

Arguments

| | |
|-----|---|
| x | output of class cquad |
| ... | further arguments passed to or from other methods |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigini (University of Ancona "Politecnica delle Marche")

| | |
|-----------|--|
| quasi_sym | <i>Resursive computation of the conditional likelihood for the Quadratic Exponential Model proposed in Bartolucci & Nigro (2010)</i> |
|-----------|--|

Description

Recursively compute the denominator of the individual conditional likelihood function for the Quadratic Exponential Model, adapted from Krailo & Pike (1984).

Usage

```
quasi_sym(eta, s, dyn=FALSE, y0=NULL)
```

Arguments

| | |
|-----|---|
| eta | individual vector of products between covariate and parameters |
| s | total score of the individual |
| dyn | TRUE if in the dynamic version; FALSE for the static version (by default) |
| y0 | Individual initial observation for dynamic models |

Value

| | |
|-----|---|
| f | value of the denominator |
| d1 | first derivative of the recursive function |
| d11 | a component of the score function |
| D2 | second derivative of the recursive function |
| D12 | a component of the Hessian matrix |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

- Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.
- Krailo, M. D., & Pike, M. C. (1984). Algorithm AS 196: conditional multivariate logistic analysis of stratified case-control studies, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **33**(1), 95-103.

| | |
|---------------|--|
| quasi_sym_equ | <i>Resursive computation of the conditional likelihood for the Modified Quadratic Exponential Model proposed in Bartolucci et al. (2018)</i> |
|---------------|--|

Description

Recursively compute the denominator of the individual conditional likelihood function for the Modified Quadratic Exponential Model recursively, adapted from Krailo & Pike (1984).

Usage

```
quasi_sym_equ(eta, s, y0=NULL)
```

Arguments

| | |
|-----|--|
| eta | individual vector of products between covariate and parameters |
| s | total score of the individual |
| y0 | Individual initial observation for dynamic models |

Value

| | |
|-----|---|
| f | value of the denominator |
| d1 | first derivative of the recursive function |
| d11 | a component of the score function |
| D2 | second derivative of the recursive function |
| D12 | a component of the Hessian matrix |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.

Bartolucci, F., Nigro, V., & Pigni, C. (2018). Testing for state dependence in binary panel data with individual covariates by a modified quadratic exponential model. *Econometric Reviews*, **37(1)**, 61-88.

Krailo, M. D., & Pike, M. C. (1984). Algorithm AS 196: conditional multivariate logistic analysis of stratified case-control studies, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **33(1)**, 95-103.

| | |
|------------------|---|
| quasi_sym_pseudo | <i>Resursive computation of pseudo conditional maximum likelihood method proposed by Bartolucci & Nigro (2012).</i> |
|------------------|---|

Description

Recursively compute the denominator of the individual conditional likelihood function for the pseudo conditional maximum likelihood method proposed by Bartolucci & Nigro (2012) recursively, adapted from Krailo & Pike (1984).

Usage

```
quasi_sym_pseudo(eta, qi, s, y0=NULL)
```

Arguments

| | |
|-----|--|
| eta | individual vector of products between covariate and parameters |
| s | total score of the individual |
| qi | Vector of quantities from first step estimation |
| y0 | Individual initial observation for dynamic models |

Value

| | |
|-----|---|
| f | value of the denominator |
| d1 | first derivative of the recursive function |
| d11 | a component of the score function |
| D2 | second derivative of the recursive function |
| D12 | a component for the Hessian matrix |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche"), Francesco Valentini (University of Ancona "Politecnica delle Marche")

References

Bartolucci, F. and Nigro, V. (2010), A dynamic model for binary panel data with unobserved heterogeneity admitting a root-n consistent conditional estimator, *Econometrica*, **78**, 719-733.

Bartolucci, F. and Nigro, V. (2012), Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data, *Journal of Econometrics*, **170**, 102-116.

Krailo, M. D., & Pike, M. C. (1984). Algorithm AS 196: conditional multivariate logistic analysis of stratified case-control studies, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **33(1)**, 95-103.

| | |
|-----------------|---|
| sim_panel_logit | <i>Simulate data from the dynamic logit model</i> |
|-----------------|---|

Description

Simulate data from the dynamic logit model given a set of covariates and a vector of parameters.

Usage

```
sim_panel_logit(id, a1, X = NULL, eta, dyn = FALSE)
```

Arguments

| | |
|-----|---|
| id | list of the reference unit of each observation |
| a1 | list of individual specific effects |
| X | corresponding matrix of covariates (optional) |
| eta | vector of parameters |
| dyn | TRUE if in the dynamic version; FALSE for the static version (by default) |

Value

| | |
|----|---|
| yv | simulated vector of binary response variables |
| pv | vector of probabilities of "success" |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigni (University of Ancona "Politecnica delle Marche")

Examples

```

# simulate data from the static logit model
n = 1000; TT = 5 # sample size, number of time occasions
id = (1:n)%x%rep(1,TT) # vector of indices
a1 = rnorm(n) # simulate alpha
X = matrix(rnorm(2*n*TT),n*TT,2) # simulate two covariates
eta1 = c(1,-1) # vector of parameters
out = sim_panel_logit(id,a1,X,eta1)
y1 = out$yv

# simulate data from the dynamic logit model
eta2 = c(1,-1,2) # vector of parameters including state dependence
out = sim_panel_logit(id,a1,X,eta2,dyn=TRUE)
y2 = out$yv

```

sq *Generate binary sequences*

Description

Generate binary sequences of a certain length and with a certain sum.

Usage

```
sq(J, s = NULL)
```

Arguments

J length of the binary sequences
s sum of the binary sequences (optional)

Value

M Matrix of binary configurations

Author(s)

Francesco Bartolucci (University of Perugia)

Examples

```

# generate all sequence of 5 binary variables
sq(5)
# generate all sequence of 5 binary variables, with sum equal 2
sq(5,2)

```

`summary.cquad`*Summary for class cquad*

Description

Summarize the output for class cquad provided by cquad_basic, cquad_equ, cquad_ext, cquad_pseudo

Usage

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

Arguments

| | |
|--------|---|
| object | output of class cquad |
| ... | further arguments passed to or from other methods |

Author(s)

Francesco Bartolucci (University of Perugia), Claudia Pigini (University of Ancona "Politecnica delle Marche")

Index

- * **conditional inference**
 - sq, [15](#)
- * **datasets**
 - data_sim, [10](#)
- * **panel data**
 - cquad, [3](#)
 - cquad-package, [2](#)
 - cquad_basic, [5](#)
 - cquad_equ, [6](#)
 - cquad_ext, [7](#)
 - cquad_pseudo, [9](#)
 - data_sim, [10](#)
 - quasi_sym, [11](#)
 - quasi_sym_equ, [12](#)
 - quasi_sym_pseudo, [13](#)
 - sim_panel_logit, [14](#)

[cquad, 3](#)
[cquad-package, 2](#)
[cquad_basic, 5](#)
[cquad_equ, 6](#)
[cquad_ext, 7](#)
[cquad_pseudo, 9](#)

[data_sim, 10](#)

[print.cquad, 11](#)

[quasi_sym, 11](#)
[quasi_sym_equ, 12](#)
[quasi_sym_pseudo, 13](#)

[sim_panel_logit, 14](#)
[sq, 15](#)
[summary.cquad, 16](#)