

# Package ‘qlspack’

July 24, 2009

**Version** 2.2

**Date** 2009-07-24

**Title** Quasi Least Square Package

**Author** Jichun Xie <jichun@mail.med.upenn.edu>, Justine Shults <jshults@mail.med.upenn.edu>

**Maintainer** Jichun Xie <jichun@mail.med.upenn.edu>

**Depends** R (>= 2.4.1), geepack (>= 1.0-12), utils

**Description** QLS is a two-stage computational approach for estimation of the correlation parameters within the framework of GEE. It helps solving parameters in mean, scale, and correlation structures for longitudinal data.

**License** GPL (>= 2)

**URL** <http://www.cceb.upenn.edu/~sratclif/QLSproject.html>

**Repository** CRAN

**Date/Publication** 2009-07-24 06:06:31

## R topics documented:

qls . . . . .	2
rat . . . . .	3
<b>Index</b>	<b>5</b>

**Description**

The `qls` function fits quasi least square estimating equations based on the `geeglm` function in the `geepack` and `cor.estimate` function in the `qlspack`. `qls` has a syntax similar to `glm` and returns an object similar to a `glm` object. An important feature of `qls`, is that an `anova` method exists for these models.

**Usage**

```
qls(formula, data, id, family = "gaussian",
    time = NULL, correlation = "ar1", std.err = "san.se")
```

**Arguments**

<code>formula</code>	The model to be fitted. The form is similar to the item documented in <code>geeglm</code> .
<code>data</code>	A data frame containing the variables in the model.
<code>id</code>	a vector which identifies the clusters. The length of 'id' should be the same as the number of observations. Data are assumed to be sorted so that observations on a cluster are contiguous rows for all entities in the formula. The 'id's for different clusters should be different, but need not to be consecutive.
<code>family</code>	A character string describing the error distribution and link function to be used in the model. There are three options: "guassian", "binomial" and "poisson". The default option is "guassian".
<code>time</code>	a vector which identifies the time in the clusters. The length of 'time' should be the same as the number of observations. This argument is used if and only if 'correlation == "markov"'.
<code>correlation</code>	a character string specifying the correlation structure. The following are permitted: "ar1", "exchangeable", "markov", "tridiagonal", "fam" and "ex.fam".
<code>std.err</code>	See corresponding documentation to <code>geeglm</code> .

**Value**

An object of type 'qlsglm'.

**Warning**

`qls` has not been thoroughly tested. Please report bugs.

**Note**

`qls` only works for complete data. Thus if there are NA's in data you can specify `data=na.omit(mydata)`.

**Author(s)**

Jichun Xie, jichun@mail.med.upenn.edu

**References**

Chaganty, N. R. 1997. An alternative approach to the analysis of longitudinal data via generalized estimating equations. *Journal of Statistical Planning and Inference* **63**: 39–54.

Xie, J. and Shults, J. 2009. Implementation of quasi-least squares With the R package qlspack. *UPenn Biostatistics Working Papers* **32**. <http://biostats.bepress.com/upennbiostat/papers/art32>

Shults, J. 1996. The analysis of unbalanced and unequally spaced longitudinal data using quasi-least squares. Ph.D. Thesis, Department of Mathematics and Statistics, Old Dominion University: Norfolk, Virginia.

Shults, J. and Chaganty, N.R. 1998. Analysis of serially correlated data using quasi-least squares. *Biometrics* **54**: 1622–1630.

Chaganty, N.R. and Shults, J. 1999. On eliminating the asymptotic bias in the quasi-least squares estimate of the correlation parameter. *Journal of Statistical Planning and Inference* **76**: 127–144.

**See Also**

[glm](#)

**Examples**

```
require(qlspack)
data(rat)
qlsfit.fam <- qls(bp ~ time + as.factor(group), data = rat, id = rat$id,
                time = rat$time, correlation = "fam")
summary(qlsfit.fam)
```

---

rat

~~ data name/kind ... ~~

---

**Description**

~~ A concise (1-5 lines) description of the dataset. ~~

**Usage**

```
data(rat)
```

**Format**

A data frame with 291 observations on the following 6 variables.

**id2** a numeric vector

**id** a numeric vector

**time** a numeric vector

**group** a numeric vector

**bp** a numeric vector

**highbp** a numeric vector

**Details**

The description of each column of the data set is as follows:

id2 = the id variable for each rat that is provided in Davis (2002).

id = a new id variable that takes value 1,2,..43 after sorting on id and group.

time = the timing of each measurement.

group = the group variable that takes value 1, 2, 3, or 4

bp = the blood pressure value.

highbp = a variable that takes value 1 if the rat's blood pressure is at least 100.

**Source**

Table 6.11 of Davis (2002)

**References**

Davis, C.(2002). *Statistical Methods for the Analysis of Repeated Measurements*.

**Examples**

```
data(rat)
## maybe str(rat) ; plot(rat) ...
```

# Index

\*Topic **datasets**

rat, 3

\*Topic **models**

qls, 1

glm, 3

qls, 1

rat, 3