

Package ‘fastqrs’

July 22, 2025

Title Fast Algorithms for Quantile Regression with Selection

Version 1.0.0

Description Fast estimation algorithms to implement the Quantile Regression with Selection estimator and the multiplicative Bootstrap for inference. This estimator can be used to estimate models that feature sample selection and heterogeneous effects in cross-sectional data. For more details, see Arellano and Bonhomme (2017) <[doi:10.3982/ECTA14030](https://doi.org/10.3982/ECTA14030)> and Pereda-Fernández (2024) <[doi:10.48550/arXiv.2402.16693](https://doi.org/10.48550/arXiv.2402.16693)>.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

Imports quantreg, copula, stats

Suggests knitr, rmarkdown, sampleSelection, ggplot2

Depends R (>= 2.10)

LazyData true

VignetteBuilder knitr

NeedsCompilation no

Author Santiago Pereda-Fernandez [aut, cre] (ORCID:
<<https://orcid.org/0000-0002-0319-1629>>)

Maintainer Santiago Pereda-Fernandez <santiagopereda@gmail.com>

Repository CRAN

Date/Publication 2025-04-16 20:10:02 UTC

Contents

.bt.results	2
.qrs.prop.fast	2
.rqr.fast	3
.rqr0.fast	4
.rqrtau.fast	5
Mroz87	5
qrs.fast	7
qrs.fast.bt	8

Index**11**

<i>.bt.results</i>	<i>bt.results</i>
--------------------	-------------------

Description

Collects the bootstrap results and yields the bootstrapped mean, standard error, and confidence intervals

Usage

```
.bt.results(x, alpha)
```

Arguments

x	= Vector with the bootstrap repetitions of an estimator
alpha	= Set of significance levels to be returned

Value

m = Bootstrapped mean
se = Bootstrapped standard error
ub = Upper bound of the bootstrapped confidence interval
lb = Lower bound of the bootstrapped confidence interval

<i>.qrs.prop.fast</i>	<i>qrs.prop.fast</i>
-----------------------	----------------------

Description

Algorithm 3: algorithm with preprocessing and quantile grid reduction for Quantile Regression with Selection (QRS); propensity score estimated previously.

Usage

```
.qrs.prop.fast(y, x, prop, w = NULL, Q1, Q2, P = 10, family, gridtheta, m)
```

Arguments

- y = Dependent variable (N x 1)
- x = Regressors matrix (N x K)
- prop = Propensity score (N x 1)
- w = Sample weights (N x 1)
- Q1 = Number of quantiles in reduced grid
- Q2 = Number of quantiles in large grid
- P = Number of evaluated values of parameter with large quantile grid
- family = Parametric copula family
- gridtheta = Grid of values for copula parameter (T x 1)
- m = Parameter to select interval of observations in top and bottom groups

Value

- beta = Estimated beta coefficients (K x Q2)
- theta = Estimated copula parameter
- objf_min = Value of objective function at the optimum
- b1 = Estimated beta coefficients for the grid of values of the copula parameter with the reduced quantile grid (K x Q1 x T)
- objf1 = Value of objective function for the grid of values of the copula parameter with the reduced quantile grid
- gridtheta2 = Grid of values for copula parameter selected during the first part of the algorithm (P x 1)
- b2 = Estimated beta coefficients for the grid of values of the copula parameter with large quantile grid (K x Q2 x P)
- objf2 = Value of objective function for the grid of values of the copula parameter with large quantile grid (P x 1)

.rqr.fast

rqr.fast

Description

Algorithm 2: algorithm with preprocessing for Rotated Quantile Regression (RQR) for a grid of quantiles and the rotation obtained with a copula.

Usage

.rqr.fast(y, x, w = NULL, G, zeta, m, initq)

Arguments

<i>y</i>	= Dependent variable (N x 1)
<i>x</i>	= Regressors matrix (N x K)
<i>w</i>	= Sample weights (N x 1)
<i>G</i>	= Copula conditional on participation (N x Q)
<i>zeta</i>	= Conservative estimate of standard error of residuals (N x 1)
<i>m</i>	= Parameter to select interval of observations in top and bottom groups
<i>initq</i>	= Initial quantile to estimate regularly and obtain preliminary values for remaining quantiles

Value

b = Estimated beta coefficients (K x Q)

.rqrb0.fast

rqrb0.fast

Description

Algorithm with preprocessing for Rotated Quantile Regression (RQR) for a grid of quantiles, the rotation obtained with a copula and initial values of the beta coefficients (used by Algorithms 3-4).

Usage

.rqrb0.fast(*y*, *x*, *w* = NULL, *G*, *zeta*, *m*, *b0*)

Arguments

<i>y</i>	= Dependent variable (N x 1)
<i>x</i>	= Regressors matrix (N x K)
<i>w</i>	= Sample weights
<i>G</i>	= Copula conditional on participation (N x Q)
<i>zeta</i>	= Conservative estimate of standard error of residuals
<i>m</i>	= Parameter to select interval of observations in top and bottom groups
<i>b0</i>	= Initial values of the beta coefficients for all quantiles (K x Q)

Value

b = Estimated beta coefficients (K x Q)

.rqrtau.fast *rqrtau.fast*

Description

Algorithm 1: algorithm with preprocessing for Rotated Quantile Regression (RQR) with initial values of the beta coefficients for a single quantile tau

Usage

.rqrtau.fast(y, x, w = NULL, tau, zeta, m, b0)

Arguments

- y = Dependent variable (N x 1)
- x = Regressors matrix (N x K)
- w = Sample weights (N x 1)
- tau = Quantile indexes rotated at individual level (N x 1)
- zeta = Conservative estimate of standard error of residuals (N x 1)
- m = Parameter to select interval of observations in top and bottom groups
- b0 = Initial values of the beta coefficients (K x 1)

Value

b = Estimated beta coefficients (K x 1)

Mroz87 *Mroz87: U.S. Women's Labor Force Participation (Example dataset)*

Description

The Mroz87 data frame contains data about 753 married women. These data are collected within the "Panel Study of Income Dynamics" (PSID). Of the 753 observations, the first 428 are for women with positive hours worked in 1975, while the remaining 325 observations are for women who did not work for pay in 1975. A more complete discussion of the data is found in Mroz (1987), Appendix 1.

Usage

Mroz87

Format

A data frame with 753 observations on the following variables:

- lfp** Dummy variable for labor-force participation.
- hours** Wife's hours of work in 1975.
- kids5** Number of children 5 years old or younger.
- kids618** Number of children 6 to 18 years old.
- age** Wife's age.
- educ** Wife's educational attainment, in years.
- wage** Wife's average hourly earnings, in 1975 dollars.
- repwage** Wife's wage reported at the time of the 1976 interview.
- hushrs** Husband's hours worked in 1975.
- husage** Husband's age.
- huseduc** Husband's educational attainment, in years.
- huswage** Husband's wage, in 1975 dollars.
- faminc** Family income, in 1975 dollars.
- mtr** Marginal tax rate facing the wife.
- motheduc** Wife's mother's educational attainment, in years.
- fatheduc** Wife's father's educational attainment, in years.
- unem** Unemployment rate in county of residence, in percentage points.
- city** Dummy variable = 1 if live in large city, else 0.
- exper** Actual years of wife's previous labor market experience.
- nwifeinc** Non-wife income.
- wifecoll** Dummy variable for wife's college attendance.
- huscoll** Dummy variable for husband's college attendance.

Source

Mroz, T. A. (1987) "The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions." *Econometrica* **55**, 765–799. PSID Staff, The Panel Study of Income Dynamics, Institute for Social Research Panel Study of Income Dynamics, University of Michigan, (For more information, visit the PSID website.).

qrs.fast *qrs.fast*

Description

Estimation of Quantile Regression with Selection (QRS) using Algorithm 3 for the estimation of the quantile and copula coefficients.

Usage

```
qrs.fast(y, x, d, z, w = NULL, Q1, Q2, P = 10, link, family, gridtheta, m = 1)
```

Arguments

y	= Dependent variable (N x 1)
x	= Regressors matrix (N x K)
d	= Participation variable (N x 1)
z	= Regressors and instruments matrix for the propensity score (N x Kz)
w	= Sample weights (N x 1)
Q1	= Number of quantiles in reduced grid
Q2	= Number of quantiles in large grid
P	= Number of evaluated values of parameter with large quantile grid
link	= Link function to compute the propensity score
family	= Parametric copula family
gridtheta	= Grid of values for copula parameter (T x 1)
m	= Parameter to select interval of observations in top and bottom groups

Value

gamma = Estimated gamma coefficients (Kz x 1)
 beta = Estimated beta coefficients (K x Q2)
 theta = Estimated copula parameter
 objf = Value of objective function at the optimum
 b1 = Estimated beta coefficients for the grid of values of the copula parameter with the reduced quantile grid (K x Q1 x T)

Examples

```

set.seed(1)

N <- 100
x <- cbind(1, 2 + runif(N))
z <- cbind(x, runif(N))
cop <- copula::normalCopula(param = -0.5, dim = 2)
copu <- copula::rCopula(N, cop)
v <- copu[,1]
u <- copu[,2]
gamma <- c(-1.5, 0.05, 2)
beta <- cbind(qnorm(u), u^0.5)
prop <- exp(z %*% gamma) / (1 + exp(z %*% gamma))
d <- as.numeric(v <= prop)
y <- d * rowSums(x * beta)
w <- matrix(1, nrow = N, ncol = 1)

Q1 <- 9
Q2 <- 19
P <- 2
m <- 1
gridtheta <- seq(from = -1, to = 0, by = .1)
link <- "probit"
family <- "Gaussian"

result <- qrs.fast(y, x[,-1], d, z[,-1], w, Q1, Q2, P, link, family, gridtheta, m)
summary(result)

```

qrs.fast.bt

qrs.fast.bt

Description

Algorithm 4: bootstrap algorithm with preprocessing and quantile grid reduction for Quantile Regression with Selection (QRS).

Usage

```

qrs.fast.bt(
  y,
  x,
  d,
  z,
  w0 = NULL,
  Q1,
  Q2,
  P = 10,
  link,

```



```

    family,
    gridtheta,
    m,
    b0,
    reps,
    alpha
)

```

Arguments

y	= Dependent variable (N x 1)
x	= Regressors matrix (N x K)
d	= Participation variable (N x 1)
z	= Regressors and instruments matrix for the propensity score (N x Kz)
w0	= Sample weights (N x 1)
Q1	= Number of quantiles in reduced grid
Q2	= Number of quantiles in large grid
P	= Number of evaluated values of parameter with large quantile grid
link	= Link function to compute the propensity score
family	= Parametric copula family
gridtheta	= Grid of values for copula parameter (T x 1)
m	= Parameter to select interval of observations in top and bottom groups
b0	= Initial values of the beta coefficients for all quantiles in the reduced quantile grid (K x Q1)
reps	= Number of bootstrap repetitions
alpha	= Significance level

Value

gammase = Bootstrapped standard error of gamma coefficients (Kz x 1)
 gammaub = Bootstrapped upper bound of confidence interval of gamma coefficients (Kz x 1)
 gammalb = Bootstrapped lower bound of confidence interval of gamma coefficients (Kz x 1)
 betase = Bootstrapped standard error of beta coefficients (K x Q)
 betaub = Bootstrapped upper bound of confidence interval of beta coefficients (K x Q)
 betalb = Bootstrapped lower bound of confidence interval of beta coefficients (K x Q)
 thetase = Bootstrapped standard error of theta coefficients (1 x 1)
 thetaub = Bootstrapped upper bound of confidence interval of theta coefficients (1 x 1)
 thetalb = Bootstrapped lower bound of confidence interval of theta coefficients (1 x 1)
 gamma = Bootstrapped estimated theta coefficients (Kz x reps)
 beta = Bootstrapped estimated beta coefficients (K x Q2 x reps)
 theta = Bootstrapped estimated copula parameter (1 x reps)
 objf = Bootstrapped value of objective function at the optimum (1 x reps)

Examples

```
set.seed(1)
N <- 100
x <- cbind(1, 2 + runif(N))
z <- cbind(x, runif(N))
cop <- copula::normalCopula(param = -0.5, dim = 2)
copu <- copula::rCopula(N, cop)
v <- copu[,1]
u <- copu[,2]
gamma <- c(-1.5, 0.05, 2)
beta <- cbind(qnorm(u), u^0.5)
prop <- exp(z %*% gamma) / (1 + exp(z %*% gamma))
d <- as.numeric(v <= prop)
y <- d * rowSums(x * beta)
w <- matrix(1, nrow = N, ncol = 1)

Q1 <- 9
Q2 <- 19
P <- 2
m <- 1
gridtheta <- seq(-1, 0, by = 0.1)
link <- "probit"
family <- "Gaussian"
reps <- 10
alpha <- 0.05

est <- qrs.fast(y, x[,-1], d, z[,-1], w, Q1, Q2, P, link, family, gridtheta, m)
bt <- qrs.fast.bt(y, x[,-1], d, z[,-1], w, Q1, Q2, P, link, family,
                 gridtheta, m, est$b1, reps, alpha)
summary(bt)
```

Index

* datasets

- Mroz87, [5](#)
- .bt.results, [2](#)
- .qrs.prop.fast, [2](#)
- .rqr.fast, [3](#)
- .rqrb0.fast, [4](#)
- .qrtau.fast, [5](#)

Mroz87, [5](#)

qrs.fast, [7](#)
qrs.fast.bt, [8](#)