

Package ‘PDMIF’

October 12, 2022

Title Fits Heterogeneous Panel Data Models

Version 0.1.0

Description Fits heterogeneous panel data models with interactive effects for linear regression, logistic, count, probit, quantile, and clustering. Based on Ando, T. and Bai, J. (2015) “A simple new test for slope homogeneity in panel data models with interactive effects” <[doi:10.1016/j.econlet.2015.09.019](https://doi.org/10.1016/j.econlet.2015.09.019)>, Ando, T. and Bai, J. (2015) “Asset Pricing with a General Multifactor Structure” <[doi:10.1093/jjfinex/nbu026](https://doi.org/10.1093/jjfinex/nbu026)>, Ando, T. and Bai, J. (2016) “Panel data models with grouped factor structure under unknown group membership” <[doi:10.1002/jae.2467](https://doi.org/10.1002/jae.2467)>, Ando, T. and Bai, J. (2017) “Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures” <[doi:10.1080/01621459.2016.1195743](https://doi.org/10.1080/01621459.2016.1195743)>, Ando, T. and Bai, J. (2020) “Quantile comovement in financial markets” <[doi:10.1080/01621459.2018.1543598](https://doi.org/10.1080/01621459.2018.1543598)>, Ando, T., Bai, J. and Li, K. (2021) “Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity” <[doi:10.1016/j.jeconom.2020.11.013](https://doi.org/10.1016/j.jeconom.2020.11.013)>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

LazyDataCompression xz

RoxygenNote 7.1.1

Imports diagonals, ncvreg, quantreg

Depends R (>= 2.10)

NeedsCompilation no

Author Tomohiro Ando [aut, cre],
Hani Fayad [aut]

Maintainer Tomohiro Ando <t.ando@mbs.edu>

Repository CRAN

Date/Publication 2022-02-16 20:20:14 UTC

R topics documented:

data1X	2
data1Y	3
data2X	3
data2Y	4
data3X	4
data3Y	5
data4LAB	5
data4X	6
data4Y	6
data5X	7
data5Y	7
data6X	8
data6Y	8
data7X	9
data7Y	9
data8Y	10
HOMTEST	10
HOMTESTGLM	11
HYPTEST	12
PDMIFCLUST	13
PDMIFCLUSTGLM	14
PDMIFCOUNT	16
PDMIFGLM	17
PDMIFLIN	18
PDMIFLING	19
PDMIFLOGIT	20
PDMIFPROBIT	21
PDMIFQUANTILE	22
PDMIFQVAR	23
Index	24

data1X	<i>A synthesized input variable dataset to fit a linear model on a panel dataset.</i>
--------	---------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to fit a linear model on a panel dataset.

Usage

data1X

Format

A data frame with 5,000 rows and 2 columns:

columns the two independent variables

rows each 100 rows represent the timeseries of each of the 50 individuals ...

data1Y	<i>A synthesized output variable dataset to fit a linear model on a panel dataset.</i>
--------	----------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a linear model on a panel dataset.

Usage

data1Y

Format

A data frame with 100 rows and 50 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data2X	<i>A synthesized input variable dataset to fit a binomial model on a panel dataset.</i>
--------	-----------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to fit a binomial model on a panel dataset.

Usage

data2X

Format

A data frame with 5,000 rows and 2 columns:

columns the two independent variables

rows each 50 rows represent the timeseries of each of the 100 individuals ...

data2Y	<i>A synthesized output variable dataset to fit a binomial model on a panel dataset.</i>
--------	------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a binomial model on a panel dataset.

Usage

data2Y

Format

A data frame with 50 rows and 100 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data3X	<i>A synthesized input variable dataset to fit a poisson model on a panel dataset.</i>
--------	----------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to fit a poisson model on a panel dataset.

Usage

data3X

Format

A data frame with 5,000 rows and 3 columns:

columns the three independent variables

rows each 50 rows represent the timeseries of each of the 100 individuals ...

data3Y	<i>A synthesized output variable dataset to fit a poisson model on a panel dataset.</i>
--------	-----------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a poisson model on a panel dataset.

Usage

data3Y

Format

A data frame with 50 rows and 100 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data4LAB	<i>A synthesized vector of memberships needed to fit a linear model on a panel dataset under known group memberships.</i>
----------	---------------------------------------------------------------------------------------------------------------------------

Description

A synthesized vector of memberships needed to fit a linear model on a panel dataset under known group memberships.

Usage

data4LAB

Format

A vector with 300 entries indicating the group membership of each individual.

data4X	<i>A synthesized input variable dataset to fit a linear model on a panel dataset under known group memberships.</i>
--------	---------------------------------------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to fit a linear model on a panel dataset under known group memberships.

Usage

data4X

Format

A data frame with 30,000 rows and 2 columns:

columns the two independent variables

rows each 100 rows represent the timeseries of each of the 300 individuals ...

data4Y	<i>A synthesized output variable dataset to fit a linear model on a panel dataset under known group memberships.</i>
--------	----------------------------------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a linear model on a panel dataset under known group memberships.

Usage

data4Y

Format

A data frame with 100 rows and 300 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data5X	<i>A synthesized input variable dataset to cluster individuals by heterogeneous panel data models with interactive effects.</i>
--------	---------------------------------------------------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to cluster individuals by heterogeneous panel data models with interactive effects.

Usage

data5X

Format

A data frame with 30,000 rows and 2 columns:

columns the two independent variables

rows each 100 rows represent the timeseries of each of the 300 individuals ...

data5Y	<i>A synthesized output variable dataset to cluster individuals by heterogeneous panel data models with interactive effects.</i>
--------	----------------------------------------------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to cluster individuals by heterogeneous panel data models with interactive effects.

Usage

data5Y

Format

A data frame with 100 rows and 300 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data6X	<i>A synthesized input variable dataset to cluster individual units by nonlinear heterogeneous panel data models with interactive effects when the group membership is unknown</i>
--------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to cluster individual units by nonlinear heterogeneous panel data models with interactive effects when the group membership is unknown

Usage

data6X

Format

A data frame with 4,500 rows and 2 columns:

columns the two independent variables

rows each 50 rows represent the timeseries of each of the 90 individuals ...

data6Y	<i>A synthesized output variable dataset to cluster individual units by nonlinear heterogeneous panel data models with interactive effects when the group membership is unknown.</i>
--------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to cluster individual units by nonlinear heterogeneous panel data models with interactive effects when the group membership is unknown.

Usage

data6Y

Format

A data frame with 50 rows and 90 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data7X	<i>A synthesized input variable dataset to fit a quantile panel data model on a panel dataset.</i>
--------	----------------------------------------------------------------------------------------------------

Description

A synthesized input variable dataset to fit a quantile panel data model on a panel dataset.

Usage

data7X

Format

A data frame with 20,000 rows and 3 columns:

columns the three independent variables

rows each 100 rows represent the timeseries of each of the 200 individuals ...

data7Y	<i>A synthesized output variable dataset to fit a quantile panel data model on a panel dataset.</i>
--------	-----------------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a quantile panel data model on a panel dataset.

Usage

data7Y

Format

A data frame with 100 rows and 200 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

data8Y	<i>A synthesized output variable dataset to fit a quantile VAR model with interactive effects and lag=2.</i>
--------	--------------------------------------------------------------------------------------------------------------

Description

A synthesized output variable dataset to fit a quantile VAR model with interactive effects and lag=2.

Usage

data8Y

Format

A data frame with 102 rows and 15 columns:

columns the individuals

rows the time points in the timeseries of each individual ...

HOMTEST	<i>HOMTEST</i>
---------	----------------

Description

This function tests homogeneity of the regression coefficients in heterogeneous panel data models with interactive effects.

Usage

HOMTEST(X, Y, Nfactors, Maxit = 100, tol = 0.001)

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- pvalue: The p-value of the homogeneity test.

References

Ando, T. and Bai, J. (2015) A simple new test for slope homogeneity in panel data models with interactive effects. *Economics Letters*, 136, 112-117.

Examples

```
fit <- HOMTEST(data1X,data1Y,2,20,0.5)
```

HOMTESTGLM

HOMTESTGLM

Description

This function tests homogeneity of the regression coefficients in heterogeneous generalized linear models with interactive effects.

Usage

```
HOMTESTGLM(X, Y, FAMILY, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
FAMILY	A description of the error distribution and link function to be used in the model just like in glm functions.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- pvalue: The p-value of the homogeneity test.

References

Ando, T. and Bai, J. (2015) A simple new test for slope homogeneity in panel data models with interactive effects. *Economics Letters*, 136, 112-117.

Examples

```
fit <- HOMTESTGLM(data2X,data2Y,binomial(link=logit),2,10,0.5)
```

HYPTEST

HYPTEST

Description

This function undergoes hypothesis testing for regression coefficients obtained from the various functions in the package.

Usage

```
HYPTEST(  
  B,  
  B0,  
  Se,  
  test = "two",  
  variables = seq(1, nrow(B)),  
  individuals = seq(1, ncol(B))  
)
```

Arguments

- | | |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| B | A dataframe of Coefficients as obtained in the output of any function in the package. |
| B0 | A dataframe of hypothetical coefficients to be evaluated in the test. (nrows should match number of variables and ncols should match number of individuals) |
| Se | A dataframe of Standard Errors as obtained in the output of any function in the package. |

test	A string to determine what kind of test to run ("two" for two-tailed, "right" for right-tailed and "left" for left-tailed).
variables	A list of variables whose coefficients are to be tested. Default is all variables in the B dataframe.
individuals	A list of individuals whose coefficients are to be tested. Default is all individuals in the B dataframe.

Value

A dataframe of p-values resulting from each individual test.

Examples

```
fit <- PDMIFLOGIT(data2X,data2Y,2,20,0.5)
HYPTTEST(fit$Coefficients,data.frame(c(0,1),c(-1,2)),fit$Se,"two",c(1,3),c(1,2))
```

PDMIFCLUST

PDMIFCLUST

Description

Under a pre-specified number of groups and the number of common factors, this function implements clustering for N individuals in the panels. Each of individuals in the group are subject to the group-specific unobserved common factors.

Usage

```
PDMIFCLUST(X, Y, NGfactors, NLfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
NGfactors	A pre-specified number of common factors across groups (see example).
NLfactors	A pre-specified number of factors in each groups (see example).
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Label: The estimated group membership for each of the individuals.
- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- GlobalFactors: The estimated common factors across groups.
- GlobalLoadings: The estimated factor loadings for the common factors.
- GroupFactors: The estimated group-specific factors.
- GroupLoadings: The estimated factor loadings for each group.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2016) Panel data models with grouped factor structure under unknown group membership *Journal of Applied Econometrics*, 31, 163-191.

Ando, T. and Bai, J. (2017) Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. *Journal of the American Statistical Association*, 112, 1182-1198.

Examples

```
fit <- PDMIFCLUST(data5X,data5Y,2,c(2,2,2),20,0.5)
```

PDMIFCLUSTGLM

PDMIFCLUSTGLM

Description

Under a pre-specified number of groups and the number of common factors, this function implements clustering for N individual units by nonlinear heterogeneous panel data models with interactive effects. Exponential family of distributions are used Each of individuals in the group are subject to the group-specific unobserved common factors.

Usage

```
PDMIFCLUSTGLM(X, Y, FAMILY, NLfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
FAMILY	A description of the error distribution and link function to be used in the model just like in glm functions.
NLfactors	A pre-specified number of factors in each groups (see example).
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Label: The estimated group membership for each of the individuals.
- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- GroupFactors: The estimated group-specific factors.
- GroupLoadings: The estimated factor loadings for each group.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2016) Panel data models with grouped factor structure under unknown group membership *Journal of Applied Econometrics*, 31, 163-191.

Ando, T. and Bai, J. (2017) Clustering huge number of financial time series: A panel data approach with high-dimensional predictors and factor structures. *Journal of the American Statistical Association*, 112, 1182-1198.

Examples

```
fit <- PDMIFCLUSTGLM(data6X,data6Y,binomial(link=logit),c(1,1,1),3,0.5)
```

PDMIFCOUNT

PDMIFCOUNT

Description

Under a known group membership, this function estimates heterogeneous poisson panel data models with interactive effects.

Usage

```
PDMIFCOUNT(X, Y, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T., Bai, J. and Li, K. (2021) Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity, *Journal of Econometrics*.

Examples

```
fit <- PDMIFCOUNT(data3X,data3Y,3,30,0.5)
```

PDMIFGLM

PDMIFGLM

Description

This function estimates heterogeneous panel data models with interactive effects through generalised linear models.

Usage

```
PDMIFGLM(X, Y, FAMILY, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
FAMILY	A description of the error distribution and link function to be used in the model just like in glm functions.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T., Bai, J. and Li, K. (2021) Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity, *Journal of Econometrics*.

Examples

```
fit <- PDMIFGLM(data2X,data2Y,binomial(link=logit),2,20,0.5)
```

PDMIFLIN

*PDMIFLIN***Description**

This function estimates heterogeneous panel data models with interactive effects. This function is similar version of PDMIFLING which accommodates a group structure.

Usage

```
PDMIFLIN(X, Y, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2015) Asset Pricing with a General Multifactor Structure Journal of Financial Econometrics, 13, 556-604.

Examples

```
fit <- PDMIFLIN(data1X,data1Y,2)
```

PDMIFLING

*PDMIFLING***Description**

Under a known group membership, this function estimates heterogeneous panel data models with interactive effects. Together with the regression coefficients, this function estimates the unobserved common factor structures both for across/within groups.

Usage

```
PDMIFLING(X, Y, Membership, NGfactors, NLfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Membership	A pre-specified group membership.
NGfactors	A pre-specified number of common factors across groups (see example).
NLfactors	A pre-specified number of factors in each groups (see example).
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- GlobalFactors: The estimated common factors across groups.
- GlobalLoadings: The estimated factor loadings for the common factors.
- GroupFactors: The estimated group-specific factors.
- GroupLoadings: The estimated factor loadings for each group.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2015) Asset Pricing with a General Multifactor Structure Journal of Financial Econometrics, 13, 556-604.

Examples

```
fit <- PDMIFLING(data4X,data4Y,data4LAB,2,c(2,2,2),30,0.1)
```

PDMIFLOGIT

PDMIFLOGIT

Description

This function estimates heterogeneous logistic panel data models with interactive effects.

Usage

```
PDMIFLOGIT(X, Y, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T., Bai, J. and Li, K. (2021) Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity, *Journal of Econometrics*.

Examples

```
fit <- PDMIFLOGIT(data2X,data2Y,2,20,0.5)
```

PDMIFPROBIT

PDMIFPROBIT

Description

This function estimates heterogeneous probit panel data models with interactive effects.

Usage

```
PDMIFPROBIT(X, Y, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated factor loadings for the common factors.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T., Bai, J. and Li, K. (2021) Bayesian and maximum likelihood analysis of large-scale panel choice models with unobserved heterogeneity, *Journal of Econometrics*.

Examples

```
fit <- PDMIFPROBIT(data2X,data2Y,2,20,0.5)
```

PDMIFQUANTILE

PDMIFQUANTILE

Description

This function estimates heterogeneous quantile panel data models with interactive effects.

Usage

```
PDMIFQUANTILE(X, Y, TAU, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

X	The (NT) times p design matrix, without an intercept where N=number of individuals, T=length of time series, p=number of explanatory variables.
Y	The T times N panel of response where N=number of individuals, T=length of time series.
TAU	A pre-specified quantile point.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated quantile point under a given tau.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2020) Quantile co-movement in financial markets Journal of the American Statistical Association.

Examples

```
fit <- PDMIFQUANTILE(data7X,data7Y,0.95,2,10,0.8)
```

PDMIFQVAR

PDMIFQVAR

Description

This function estimates heterogeneous quantile panel data VAR models with interactive effects.

Usage

```
PDMIFQVAR(Y, LAG, TAU, Nfactors, Maxit = 100, tol = 0.001)
```

Arguments

Y	The T times N panel of response where N=number of individuals, T=length of time series.
LAG	The number of lags from y_{t-1} to y_{t-LAG} used in the VAR.
TAU	A pre-specified quantile point.
Nfactors	A pre-specified number of common factors.
Maxit	A maximum number of iterations in optimization. Default is 100.
tol	Tolerance level of convergence. Default is 0.001.

Value

A list with the following components:

- Coefficients: The estimated heterogeneous coefficients.
- Lower05: Lower end (5%) of the 90% confidence interval of the regression coefficients.
- Upper95: Upper end (95%) of the 90% confidence interval of the regression coefficients.
- Factors: The estimated common factors across groups.
- Loadings: The estimated quantile point under a given tau.
- Predict: The conditional expectation of response variable.
- pval: p-value for testing hypothesis on heterogeneous coefficients.
- Se: Standard error of the estimated regression coefficients.

References

Ando, T. and Bai, J. (2020) Quantile co-movement in financial markets Journal of the American Statistical Association.

Examples

```
fit <- PDMIFQVAR(data8Y,2,0.1,2,5,0.8)
```

Index

* datasets

data1X, 2
data1Y, 3
data2X, 3
data2Y, 4
data3X, 4
data3Y, 5
data4LAB, 5
data4X, 6
data4Y, 6
data5X, 7
data5Y, 7
data6X, 8
data6Y, 8
data7X, 9
data7Y, 9
data8Y, 10

data1X, 2
data1Y, 3
data2X, 3
data2Y, 4
data3X, 4
data3Y, 5
data4LAB, 5
data4X, 6
data4Y, 6
data5X, 7
data5Y, 7
data6X, 8
data6Y, 8
data7X, 9
data7Y, 9
data8Y, 10

HOMTEST, 10
HOMTESTGLM, 11
HYPTTEST, 12

PDMIFCLUST, 13

PDMIFCLUSTGLM, 14
PDMIFCOUNT, 16
PDMIFGLM, 17
PDMIFLIN, 18
PDMIFLING, 19
PDMIFLOGIT, 20
PDMIFPROBIT, 21
PDMIFQUANTILE, 22
PDMIFQVAR, 23