

Package ‘fbi’

October 28, 2019

Type Package

Title Factor-Based Imputation and FRED-MD Data Set

Version 0.1.0

Date 2019-10-11

Author Yankang (Bennie) Chen [aut, cre]
Serena Ng [aut],
Jushan Bai [aut]

Maintainer Yankang (Bennie) Chen <yankang.chen@columbia.edu>

Description The fbi package contains functions to estimate factor models and impute missing data based on factor models. It also includes functions to load and prepare the FRED-MD data set.

Depends R (>= 3.5.0)

Imports readr, pracma

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

Suggests knitr, rmarkdown

VignetteBuilder knitr

R topics documented:

fbi-package	2
apc	2
describe	3
fredmd	4
fredmd_description	5
rm_outliers.fredmd	5
rpca	6
se.rpca	7
tw_apc	8

Index	9
--------------	----------

fbi-package

*Factor-Based Imputation and FRED-MD Data Set***Description**

The fbi package contains functions to estimate factor models and impute missing data based on factor models. It also includes functions to load and prepare the FRED-MD data set.

Details

See `vignette("factor_fred", package = "fbi")` for an example using the FRED-MD dataset (<https://research.stlouisfed.org/econ/mccracken/fred-databases/>).

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

Serena Ng <serena.ng@columbia.edu>

Jushan Bai <jushan.bai@columbia.edu>

References

Jushan Bai and Serena Ng (2002), *Determining the number of factors in approximate factor models*. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00273>

Jushan Bai and Serena Ng (2017), *Rank regularized estimation of approximate factor models*. <https://www.sciencedirect.com/science/article/pii/S0304407619300764>

apc

*Factor Model of Balanced Panel Data***Description**

apc estimates the factor model of a given balanced panel data.

Usage

```
apc(X, r)
```

Arguments

`X` a matrix of size T by N.
`r` integer, indicating the maximum number of factors.

Value

a list of elements:

`Fhat`

`Lamhat`

`d`

`d0`

`ehat`

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

Serena Ng <serena.ng@columbia.edu>

Jushan Bai <jushan.bai@columbia.edu>

References

Jushan Bai and Serena Ng (2019), *Matrix Completion, Counterfactuals, and Factor Analysis of Missing Data*. <https://arxiv.org/abs/1910.06677>

describe	<i>Describe selected variables in the FRED-MD Data Set</i>
----------	--

Description

describe provides a description of the selected variables in the FRED-MD data set.

Usage

```
describe(varname, name.only = TRUE, verbose = FALSE)
```

Arguments

varname	string or a vector strings of the format "X1" to "X135".
name.only	logical. If TRUE, return a dataframe with variable names and types of transformation only; if FALSE, return a dataframe with more details.
verbose	logical, indicating whether or not descriptions should be printed.

Value

a vector of variable names, or a data frame with detailed descriptions.

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

References

Michael W. McCracken and Serena Ng (2015), *FRED-MD Updated Appendix*. https://s3.amazonaws.com/files.fred.stlouisfed.org/fred-md/Appendix_Tables_Update.pdf

Examples

```
varnames <- describe(c("X32", "X56"), name.only = TRUE, verbose = FALSE)
```

`fredmd`*Loading FRED-MD Data Set*

Description

`fredmd` loads the official FRED-MD data set and provides a few tools to manipulate the data set.

Usage

```
fredmd(date_start = NULL, date_end = NULL, transform = TRUE,  
       local = FALSE)
```

Arguments

<code>date_start</code>	Date or NULL, the start date (included) of the data selection. If NULL, select till the latest data available.
<code>date_end</code>	Date or NULL, the end date (included) of the data selection. If NULL, select up to the earliest data available.
<code>transform</code>	logical, indicating Whether or not the FRED-MD data set should be transformed according to the transformation code.
<code>local</code>	logical, indicating Whether or not the FRED-MD data set should be loaded from the local files or downloaded online

Value

a subset of the (transformed) FRED-MD data of class `fredmd`.

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

References

Michael W. McCracken and Serena Ng (2015), *FRED-MD and FRED-QD: Monthly and Quarterly Databases for Macroeconomic Research*. <https://research.stlouisfed.org/econ/mccracken/fred-databases/>

Examples

```
data <- fredmd(date_start = NULL, date_end = NULL, transform = TRUE, local = FALSE)
```

fredmd_description	<i>FRED-MD Data Set Description</i>
--------------------	-------------------------------------

Description

A description of the FRED-MD data set.

Usage

```
data(fredmd_description)
```

Format

A data frame with 135 rows and 9 variables. The variables are as follows:

id series ID number

tcode type of transformation

fred variable name used in the FRED-MD data set

description description of the series

gsi variable name used in the Global Insights Basic Economics Database (GSI)

gsi:description description of the series in GSI

group group of the series

edited logical, indicating if the data has been edited

varname "X" + id

Source

The fredmd_description data were obtained from Michael W. McCracken and Serena Ng (2015), *FRED-MD Updated Appendix*. https://s3.amazonaws.com/files.fred.stlouisfed.org/fred-md/Appendix_Tables_Update.pdf

rm_outliers.fredmd	<i>Remove outliers of the FRED-MD Data Set</i>
--------------------	--

Description

rm_outliers.fredmd removes outliers of the FRED-MD data set produced by the [fredmd](#) function.

Usage

```
rm_outliers.fredmd(object)
```

Arguments

object an object of class [fredmd](#).

Value

FRED-MD data of class fredmd with outliers removed.

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

References

Michael W. McCracken and Serena Ng (2015), *FRED-MD and FRED-QD: Monthly and Quarterly Databases for Macroeconomic Research*. <https://research.stlouisfed.org/econ/mccracken/fred-databases/>

Examples

```
data <- fredmd(date_start = NULL, date_end = NULL, transform = TRUE)
newdata <- rm_outlier.fredmd(data)
```

rpca

Estimation of Approximate Factor Models

Description

rpca estimates the approximate factor models of the given matrix.

Usage

```
rpca(X, kmax, standardize = FALSE, tau = 0)
```

Arguments

X	a matrix of size T by N.
kmax	integer, indicating the maximum number of factors.
standardize	logical, indicating Whether or not X should be centered and scaled.
tau	numeric, specifying the parameter in the rank-regularized estimation. If tau = 0, then rank regularization is not used.

Value

a list of elements:

```
X
kmax
standardize
tau
ic2
pc2k
pc20
Fhat
Lamhat
Chat
Sigma
```

IC2
 PC2k
 PC20
 fhat
 lamhat
 d
 d0

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>
 Serena Ng <serena.ng@columbia.edu>
 Jushan Bai <jushan.bai@columbia.edu>

References

Jushan Bai and Serena Ng (2002), *Determining the number of factors in approximate factor models*.
<https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00273>
 Jushan Bai and Serena Ng (2017), *Rank regularized estimation of approximate factor models*.
<https://www.sciencedirect.com/science/article/pii/S0304407619300764>

se.rpca	<i>Standard Error of C^{hat}</i>
---------	--

Description

se.rpca produces the estimated standard error of C^{hat} produced by the `rpca` function.

Usage

```
se.rpca(object, xpoints, qq)
```

Arguments

object	an object of class <code>rpca</code> .
xpoints	placeholder.
qq	placeholder.

Value

standard error of C^{hat}

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>
 Serena Ng <serena.ng@columbia.edu>
 Jushan Bai <jushan.bai@columbia.edu>

References

Jushan Bai and Serena Ng (2002), *Determining the number of factors in approximate factor models*. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0262.00273>

Jushan Bai and Serena Ng (2017), *Rank regularized estimation of approximate factor models*. <https://www.sciencedirect.com/science/article/pii/S0304407619300764>

tw_apc

Tall-Wide Imputation of Missing Value in Panel Data

Description

tw_apc imputates the missing values in a given panel data.

Usage

```
tw_apc(X1, r1, center = FALSE, standardize = FALSE)
```

Arguments

X1	a matrix of size T by N.
r1	integer, indicating the maximum number of factors.
center	logical, indicating Whether or not X1 should be demeaned
standardize	logical, indicating Whether or not X1 should be scaled.

Value

a list of elements:

Fhat
Lamhat
Chat
data

Author(s)

Yankang (Bennie) Chen <yankang.chen@columbia.edu>

Serena Ng <serena.ng@columbia.edu>

Jushan Bai <jushan.bai@columbia.edu>

References

Jushan Bai and Serena Ng (2019), *Matrix Completion, Counterfactuals, and Factor Analysis of Missing Data*. <https://arxiv.org/abs/1910.06677>

Index

*Topic **datasets**

fredmd_description, [5](#)

apc, [2](#)

describe, [3](#)

fbi (fbi-package), [2](#)

fbi-package, [2](#)

fredmd, [4](#), [5](#)

fredmd_description, [5](#)

rm_outliers.fredmd, [5](#)

rpca, [6](#), [7](#)

se.rpca, [7](#)

tw_apc, [8](#)