Package 'Largevars'

May 7, 2026

Type Package

Title Testing Large VARs for the Presence of Cointegration

Version 1.0.3

Maintainer Eszter Kiss <ekiss2803@gmail.com>

Description Conducts a cointegration test for high-dimensional vector autoregressions (VARs) of order k based on the large N,T asymptotics of Bykhovskaya and Gorin, 2022 (<doi:10.48550/arXiv.2202.07150>). The implemented test is a modification of the Johansen likelihood ratio test. In the absence of cointegration the test converges to the partial sum of the Airy-1 point process. This package contains simulated quantiles of the first ten partial sums of the Airy-1 point process that are precise up to the first three digits.

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

Depends R (>= 3.5.0)

Imports methods, graphics, stats, utils

Suggests testthat (>= 3.0.0), tibble (>= 3.0.0), data.table (>= 1.14.0), readr (>= 2.1.0)

Config/testthat/edition 3

License MIT + file LICENSE

URL https://github.com/eszter-kiss/Largevars

NeedsCompilation no

Author Anna Bykhovskaya [aut],
Vadim Gorin [aut],
Eszter Kiss [cre, aut]

Repository CRAN

Date/Publication 2025-05-19 02:10:02 UTC

Contents

`largevar` *Cointegration test for settings of large N and T*

Description

Runs the Bykhovskaya-Gorin test for cointegration. Paper can be found at: https://doi.org/10.48550/arXiv.2202.07150

Usage

```
largevar(
  data = NULL,
  k = 1,
  r = 1,
  fin_sample_corr = FALSE,
  plot_output = TRUE,
  significance_level = 0.05
)
```

Arguments

`data`	A numeric matrix where the columns contain individual time series that will be examined for the presence of cointegrating relationships.
`k`	The number of lags that we wish to employ in the vector autoregression. The default value is k = 1.
`r`	The number of largest eigenvalues used in the test. The default value is r = 1.
`fin_sample_corr`	A boolean variable indicating whether we wish to employ finite sample correction on our test statistic. The default value is fin_sample_corr = FALSE.
`plot_output`	A boolean variable indicating whether we wish to generate a plot of the empirical distribution of eigenvalues. The default value plot_output = TRUE.
`significance_level`	Specify the significance level at which the decision about H0 should be made. The default value is significance_level = 0.05.

Value

A list that contains the test statistic, a table with theoretical quantiles presented for r=1 to r=10, and the decision about H0 at the significance level specified by the user.

Examples

```
largevar(
  data = matrix(rnorm(60, mean = 0.05, sd = 0.01), 20, 3),
  k = 1,
  r = 1,
  fin_sample_corr = FALSE,
  plot_output = FALSE,
  significance_level = 0.05
)
```

---

percentiles *Quantiles for the limiting distribution of the test*

Description

A data frame containing the simulated quantiles for the test statistic used in the largevar function. More details about how these simulations were conducted can be found in Section 4 of the vignette.

Format

A data frame with 99 rows and 11 variables:

Source

Calculated through own simulations (see details in vignette).

---

quantile_tables *Creates the quantile table output for largevar function*

Description

Outputs the quantile tables from the package's corresponding vignette.

Usage

```
quantile_tables(r = 1)
```

Arguments

r	Which partial sum the quantile table should be returned for. (Only r<=10 is available.) Default is r=1.

Value

A numeric matrix.

Examples

```
quantile_tables(r=3)
```

`sim_function` *Empirical p-value for cointegration test*

Description

Runs a simulation on the H0 for the Bykhovskaya-Gorin test for cointegration and returns an empirical p-value. Paper can be found at: https://doi.org/10.48550/arXiv.2202.07150

Usage

```
sim_function(
  N = NULL,
  tau = NULL,
  stat_value = NULL,
  k = 1,
  r = 1,
  fin_sample_corr = FALSE,
  sim_num = 1000,
  seed = NULL
)
```

Arguments

`N`	The number of time series used in simulations.
`tau`	The length of the time series used in simulations.
`stat_value`	The test statistic value for which the p-value is calculated.
`k`	The number of lags that we wish to employ in the vector autoregression. The default value is k = 1.
`r`	The number of largest eigenvalues used in the test. The default value is r = 1.
`fin_sample_corr`	A boolean variable indicating whether we wish to employ finite sample correction on our test statistics. The default value is fin_sample_corr = FALSE.
`sim_num`	The number of simulations that the function conducts for H0. The default value is sim_num = 1000.
`seed`	The random seed that a user can set for replicable simulation results. The default value is seed = NULL.

Value

A list that contains the simulation values, the empirical percentage (realizations larger than the test statistic provided by the user) and a histogram.

Examples

```
sim_function(N=90, tau=501, stat_value=-0.27,k=1,r=1,sim_num=30, seed = 0)
```

s_p100_price *Stock price data for example in vignette*

Description

A data frame containing weekly S&P100 prices over ten years: 01.01.2010 - 01.01.2020, The S&P100 includes 101 leading U.S. stocks of which 92 were collected here.

Format

A data frame with 522 rows and 93 variables:

Source

Refer to the data source used in: A. Bykhovskaya and V. Gorin. Cointegration in large vars. Annals of Statistics, 2022.

Index