Package: BTtest (via r-universe)

October 26, 2026

Title Estimate the Number of Factors in Large Nonstationary Datasets

Version 0.10.3

Maintainer Paul Haimerl <p.haimerl@student.maastrichtuniversity.nl>

Description Large panel data sets are often subject to common trends. However, it can be difficult to determine the exact number of these common factors and analyse their properties. The package implements the Barigozzi and Trapani (2022) <doi:10.1080/07350015.2021.1901719> test, which not only provides an efficient way of estimating the number of common factors in large nonstationary panel data sets, but also gives further insights on factor classes. The routine identifies the existence of (i) a factor subject to a linear trend, (ii) the number of zero-mean I(1) and (iii) zero-mean I(0) factors. Furthermore, the package includes the Integrated Panel Criteria by Bai (2004) <doi:10.1016/j.jeconom.2003.10.022> that provide a complementary measure for the number of factors.

License GPL (>= 3)

Encoding UTF-8

BugReports https://github.com/Paul-Haimerl/BTtest/issues

RoxygenNote 7.3.1

URL https://github.com/Paul-Haimerl/BTtest

Imports Rcpp

LinkingTo Rcpp, RcppArmadillo

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation yes

Author Paul Haimerl [aut, cre] (<https://orcid.org/0000-0003-3198-8317>)

Depends R (>= 3.5.0)

Repository https://cran.r-universe.dev

Date/Publication 2024-08-29 02:51:06 UTC

RemoteUrl https://github.com/cran/BTtest

RemoteRef HEAD

RemoteSha 3c37ebecb414e4cddb902717d2bff733c09e446f

Contents

BaiIPC 2
BTtest 3
sim_DGP 4

Index **6**

`BaiIPC` *Bai (2004) IPC*

Description

Calculates the Integrated Panel Criteria (*IPC*) to estimate the total number of common trends in a nonstationary panel as proposed by Bai (2004).

Usage

```
BaiIPC(X, r_max = 10)
```

Arguments

`X`	a $T \times N$ numerical `matrix` or `data.frame` of observations.
`r_max`	the maximum number of factors to consider. Default is 10.

Details

For further details on the three criteria and their respective differences, I refer to Bai (2004, sec. 3).

Value

A vector of the estimated number of factors for each of the three criteria.

Author(s)

Paul Haimerl

References

Bai, J. (2004). Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, 122(1), 137-183. doi:10.1016/j.jeconom.2003.10.022

Examples

```
# Simulate a nonstationary panel
X <- sim_DGP(N = 100, n_Periods = 200)

# Obtain the estimated number of common factors pre criterion
BaiIPC(X = X, r_max = 10)
```

`BTtest` *Barigozzi & Trapani (2022) Test*

Description

Runs the testing routine proposed in Barigozzi & Trapani (2022) to estimate the number and types of common trends in a nonstationary panel. The method can identify the existence of a common factor subject to a linear trend, as well as the number of zero-mean $I(1)$ and zero-mean $I(0)$ factors.

Usage

```
BTtest(X, r_max = 10, alpha = 0.05, BT1 = TRUE, R = NULL)
```

Arguments

`X`	a $T \times N$ numerical `matrix` or `data.frame` of observations.
`r_max`	the maximum number of factors to consider. Default is 10. Note that changing `r_max` does not alter the test result for any individual `r`.
`alpha`	the significance level. Default is 0.05.
`BT1`	logical. If `TRUE`, a less conservative eigenvalue rescaling scheme is used. In small samples, `BT1 = FALSE` will result in fewer estimated factors. Default is `TRUE`.
`R`	the number of draws from an *i.i.d.* standard normal random variable that constructs the randomized test statistic. If `R = NULL`, the heuristic `R`$= 2N$ is used when testing for a spike in the largest eigenvalue and else `R`$= max(100, floor(N/3))$, where N is the number of columns in `X`. See Barigozzi & Trapani (2022, sec. 5). We recommend the default `NULL`.

Details

For details on the testing procedure I refer to Barigozzi & Trapani (2022, sec. 4).

Value

A vector with the estimated number of (i) factors with a linear trend (r_1), (ii) zero-mean $I(1)$ factors (r_2) and (ii) zero-mean $I(0)$ factors (r_3).

Author(s)

Paul Haimerl

References

Barigozzi, M., & Trapani, L. (2022). Testing for common trends in nonstationary large datasets. *Journal of Business & Economic Statistics*, 40(3), 1107-1122. doi:10.1080/07350015.2021.1901719

Examples

```
# Simulate a nonstationary panel
X <- sim_DGP(N = 100, n_Periods = 200)

# Obtain the estimated number of factors (i) with a linear trend (r_1), (ii) zero-mean I(1) (r_2)
# and (iii) zero-mean I(0) (r_3)
BTtest(X = X, r_max = 10, alpha = 0.05, BT1 = TRUE)
```

`sim_DGP` *Simulate a Nonstationary Panel With Common Trends*

Description

Simulate a nonstationary panel as laid out in Barigozzi & Trapani (2022, sec. 5).

Usage

```
sim_DGP(
  N = 100,
  n_Periods = 200,
  drift = TRUE,
  drift_I1 = TRUE,
  r_I1 = 2,
  r_I0 = 1,
  return_factor = FALSE
)
```

Arguments

`N`	the number of cross-sectional units.
`n_Periods`	the number of simulated time periods.
`drift`	logical. If `TRUE`, a linear trend is included (corresponding to both *d_1 = 1* and *r_1 = 1*).
`drift_I1`	logical. If `TRUE`, an *I(1)* factor moves around the linear trend. Else an *I(0)* factor (corresponding to *d_2 = 1*).
`r_I1`	the total number of non zero-mean *I(1)* factors (corresponding to *r_2 + r_1 * d_2*).
`r_I0`	the total number of non zero-mean *I(0)* factors (corresponding to *r_3 + r_1 * (1 - d_2)*).
`return_factor`	logical. If `TRUE`, the factor matrix is returned. Else the simulated observations. Default is `FALSE`.

Details

For further details on the construction of the DGP, see Barigozzi & Trapani (2022, sec. 5).

Value

A (T x N) matrix of simulated observations. If `return_factor = TRUE`, a (N x r) matrix of factors.

Author(s)

Paul Haimerl

References

Barigozzi, M., & Trapani, L. (2022). Testing for common trends in nonstationary large datasets. *Journal of Business & Economic Statistics*, 40(3), 1107-1122. doi:10.1080/07350015.2021.1901719

Examples

```
# Simulate a panel containing a factor with a linear drift (r_1 = d_1 = 1) and I(1) process
# (d_2 = 1), one zero-mean I(1) factor (r_2 = 1) and two zero-mean I(0) factors (r_3 = 2)
X <- sim_DGP(N = 100, n_Periods = 200, drift = TRUE, drift_I1 = TRUE, r_I1 = 2, r_I0 = 2)

# Simulate a panel containing only 3 common zero-mean I(0) factor (r_1 = 0, r_2 = 0, r_3 = 3)
X <- sim_DGP(N = 100, n_Periods = 200, drift = FALSE, drift_I1 = TRUE, r_I1 = 0, r_I0 = 3)
```

Index

BaiIPC, 2
BTtest, 3

sim_DGP, 4