Package 'statioVAR'

August 20, 2025

Type Package

Title Trend Removal for Vector Autoregressive Workflows

Version 0.1.3

Description Detrending multivariate time-series to approximate stationarity when dealing with intensive longitudinal data, prior to Vector Autoregressive (VAR) or multilevel-VAR estimation. Classical VAR assumes weak stationarity (constant first two moments), and deterministic trends inflate spurious autocorrelation, biasing Granger-causality and impulse-response analyses. All functions operate on raw panel data and write detrended columns back to the data set, but differ in the level at which the trend is estimated. See, for instance, Wang & Maxwell (2015) <doi:10.1037/met0000030>; Burger et al. (2022) <doi:10.4324/9781003111238-13>; Epskamp et al. (2018) <doi:10.1177/2167702617744325>.

URL https://github.com/g-corbelli/statioVAR

BugReports https://github.com/g-corbelli/statioVAR/issues

License GPL-3

Encoding UTF-8

Imports dplyr, rlang, stats

Suggests shiny, testthat (>= 3.0.0), knitr, rmarkdown

Language en-US

NeedsCompilation no

Maintainer Giuseppe Corbelli <giuseppe.corbelli@uniroma1.it>

Config/testthat/edition 3

RoxygenNote 7.3.2

Author Giuseppe Corbelli [aut, cre] (ORCID:
<https://orcid.org/0000-0002-2864-3548>)

Repository CRAN

Date/Publication 2025-08-20 07:10:02 UTC

Contents

`detrender` *Within-unit linear detrending for multilevel VAR analysis*

Description

Remove unit-specific linear trends from panel data to approximate stationarity, preparing inputs for multilevel Vector Autoregressive (VAR) modeling (among others). For each unit (subject) and each selected variable, a linear regression of the variable on the time index is tested at significance level `alpha`; if the slope is significant, the fitted trend is subtracted and the mean of the unit is re-added, to produce detrended series while preserving between-unit information.

Caution: models with lagged outcomes and per-unit intercepts (fixed or random) are prone to Nickell-type bias when there are fewer than 10 time points (T) per unit; detrending does not remove it. `T >= 10` is recommended (Nickell, 1981; Judson & Owen, 1999). For VAR(1) with an intercept and linear trend, a minimum of `K + 4` time points per unit (where `K` is the number of detrended series) is required to maintain positive residual degrees of freedom (Lütkepohl, 2005).

Usage

```
detrender(
  df,
  id_var,
  time_var,
  vars_to_detrend,
  alpha = 0.05,
  min_obs = 3
)
```

Arguments

`df`	Data frame or tibble (long format).
`id_var`	Character string. Unit (subject) identifier column (required).
`time_var`	Character string. Numeric time index column (required).
`vars_to_detrend`	Character vector. Column names to detrend within each unit (subject) (required).
`alpha`	Numeric in (0,1). Significance threshold for retaining a non-zero time slope (default: 0.05).
`min_obs`	Integer >2. Minimum observations per unit-variable to attempt detrending (default: 3).

Value

A named list with:

`df` Tibble. The original dataset with additional detrended columns.

`n_units` Integer. Number of unique units (subjects) processed.

`total_trends` Integer. Total number of individual trends removed across all variables.

`summary` Tibble. Number of removed linear trends per variable, with columns `variable` and `removed_trends`.

References

Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. Economics letters, 65(1), 9-15. doi:10.1016/s01651765(99)001305

Lütkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. Springer Berlin Heidelberg. doi:10.1007/9783540277521

Nickell, S. (1981). Biases in dynamic models with fixed effects. Econometrica: Journal of the econometric society, 1417-1426. doi:10.2307/1911408

Examples

```
df_example <- data.frame(
id = rep(1:2, each = 5),
time = rep(1:5, 2),
x = rep(1:5, 2) + rnorm(10)
)
res <- statioVAR::detrender(
df = df_example,
id_var = "id",
time_var = "time",
vars_to_detrend = "x",
alpha = 0.05,
min_obs = 3
)
res$df[7:9,]
res$n_units
res$total_trends
res$summary
```

`pooled` *Pooled polynomial detrending for multivariate panel data*

Description

Remove study-wide polynomial trend (up to cubic) plus optional cyclic effects from multivariate panel data by fitting a single OLS model on the pooled series. Trend terms up to the chosen degree are estimated; those whose two-sided t-tests are significant at `alpha` are retained, non-significant components are set to 0, and the resulting fitted values are subtracted from every observation of the raw series.

Usage

```
pooled(
  df,
  id_var,
  time_var = NULL,
  vars_to_detrend,
  poly_order = 1,
  cyc_vars = NULL,
  alpha = 0.05,
  miss_thresh = 0.30
)
```

Arguments

df	Data frame or tibble (long format).
id_var	Character string. Unit (subject) identifier column (required).
time_var	Character string. Numeric time index column (if NULL, then cyc_vars must be specified). If NULL, no polynomial time terms are included.
vars_to_detrend	Character vector. Column names to detrend (required).
poly_order	Integer in {1,2,3}. Maximum degree of the polynomial time trend tested (default: 1): • 1 = linear only, • 2 = linear + quadratic, • 3 = linear + quadratic + cubic.
cyc_vars	Character vector. Column names (e.g. "weekend") for categorical cyclicity variables (if NULL, then time_var must be specified).
alpha	Numeric in (0,1). Significance threshold for retaining polynomial terms (default 0.05).
miss_thresh	Numeric in (0,1). Maximum allowed proportion of missing data per variable (default: 0.30).

Value

A named list with:

df Tibble with added <var>_detrended columns.

coef_tables Named list of coefficient tables (one per variable), with columns predictor, estimate, Std. Error, t, p, and a logical flag kept.

formula_str Character string of the fitted model formula.

n_units Integer: number of unique units (subjects).

Examples

```
dat <- data.frame(
id = rep(1:3, each=5),
time = rep(1:5, 3),
cyc = rep(c("A","B"), length.out=15),
y1 = rnorm(15, sd = 0.5) + seq(1,15)*1.0
)
res <- statioVAR::pooled(
df = dat,
id_var = "id",
time_var = "time",
vars_to_detrend = "y1",
poly_order = 2,
cyc_vars = "cyc",
alpha = 0.05,
miss_thresh = 0.30
)
```

`statioVAR` *Trend Removal for Vector Autoregressive Workflows*

Description

Detrending multivariate time series to approximate stationarity in intensive longitudinal data, prior to vector autoregressive (VAR) or multilevel VAR estimation. Classical VAR assumes weak stationarity (i.e., constant mean, and autocovariances that depend only on lag), and deterministic trends can induce spurious autocorrelation, distorting Granger causality and impulse-response analyses. All functions operate on raw panel data and write detrended columns back to the data set, but differ in the level at which the trend is estimated.

Details

The functions are:

- `detrender`: within-unit linear detrending, which fits and removes a separate linear trend for each unit (subject) on each selected variable.
- `pooled`: pooled polynomial detrending, which fits and removes a global polynomial trend (up to cubic) and optional cyclic effects across all units (subjects).

Note

The development of this package was inspired by, and is deeply indebted to, the works of Eiko Fried, Jonas Haslbeck, Sasha Epskamp, Ria Hoekstra and Alessandra Mansueto, among others. This software is provided 'as is', without any express or implied warranties of accuracy or reliability. For suggestions or to report any issue, please contact the author.

Author(s)

Giuseppe Corbelli (<giuseppe.corbelli@uniroma1.it>)

See Also

Useful links:

- https://github.com/g-corbelli/statioVAR
- Report bugs at https://github.com/g-corbelli/statioVAR/issues

Index