Package: stagsynth (via r-universe)

May 29, 2026

Type Package

Title Staggered Synthetic Control Estimation and Inference

Version 0.1.0

Description Implements the Staggered Synthetic Control (SSC) method for estimating treatment effects in panel data with staggered adoption, as proposed by Cao, Lu, and Wu (2020) <doi:10.48550/arXiv.1912.06320>. Constructs synthetic control weights via constrained quadratic programming, estimates heterogeneous treatment effects and event-time average treatment effects on the treated (ATT), and provides placebo-in-time confidence intervals and p-values.

License GPL (>= 3)

Encoding UTF-8

Depends R (>= 3.5.0)

Imports quadprog

Suggests ggplot2, testthat (>= 3.0.0)

Config/testthat/edition 3

RoxygenNote 7.3.3

NeedsCompilation no

Author Zhanchao Fu [aut, cre], Jianfei Cao [aut]

Maintainer Zhanchao Fu <fuzhanchao2001@gmail.com>

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

Date/Publication 2026-04-28 20:46:43 UTC

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

RemoteRef HEAD

RemoteSha 08c3213cc84be7959e1832a079904572de586ce8

Contents

`stagsynth-package` *stagsynth: Staggered Synthetic Control Estimation and Inference*

Description

Implements the Staggered Synthetic Control (SSC) method of Cao, Lu, and Wu (2020) for estimating treatment effects in panel data with staggered adoption.

Main function

`ssc`: Estimate event-time ATT, overall ATT, and placebo-in-time confidence intervals.

Utilities

- `panel_to_matrices`: Convert long-format panel data to the $N \times T$ matrices expected by `ssc()`.
- `ssc_min_eigenvalue`: Check the design matrix invertibility condition.
- `synthetic_control`: Estimate SC weights for a single treated unit.
- `synthetic_control_batch`: Estimate SC weights for all units.

Author(s)

Maintainer: Zhanchao Fu <fu.zhan@northeastern.edu>

Authors:

- Jianfei Cao <jianfeicao7@gmail.com>

`panel_to_matrices` *Convert Long-Format Panel Data to Matrices*

Description

Transform a data frame in long format (one row per unit-period) into the $N \times T$ matrices `Y` and `D` expected by `ssc`.

Usage

```
panel_to_matrices(data, unit, time, outcome, treatment)
```

Arguments

`data`	A data frame.
`unit`	Character: name of the unit identifier column.
`time`	Character: name of the time period column.
`outcome`	Character: name of the outcome variable column.
`treatment`	Character: name of the treatment indicator column (must be 0/1).

Value

A list with components

Y Numeric $N \times T$ outcome matrix.

D Numeric $N \times T$ treatment matrix.

units Sorted vector of unique unit identifiers.

times Sorted vector of unique time periods.

Examples

```
df <- data.frame(
  id   = rep(1:4, each = 6),
  time = rep(1:6, times = 4),
  Y    = rnorm(24),
  D    = c(rep(0, 12), rep(c(0,0,0,1,1,1), 2))
)
mat <- panel_to_matrices(df, unit = "id", time = "time",
                         outcome = "Y", treatment = "D")
```

plot.ssc *Plot Event-Time ATT from SSC Estimation*

Description

Plot Event-Time ATT from SSC Estimation

Usage

```
## S3 method for class 'ssc'
plot(
  x,
  main = "Event-time ATT (SSC)",
  xlab = "Event time",
  ylab = "ATT estimate",
  ci = !anyNA(x$ci_lower_event),
  ...
)
```

Arguments

x	An object of class "ssc".
main	Title string.
xlab, ylab	Axis labels.
ci	Logical: draw the confidence band? Default TRUE if inference is available.
...	Additional arguments (currently unused).

Value

A ggplot object (invisibly) if **ggplot2** is available; otherwise a base-R plot is drawn and NULL is returned invisibly.

Examples

```
set.seed(1)
N <- 10; Ttot <- 8
Y <- matrix(rnorm(N * Ttot), N, Ttot)
D <- matrix(0L, N, Ttot)
D[1:3, 5:Ttot] <- 1L   # units 1-3 treated from period 5
fit <- ssc(Y, D, S = 2, alpha = 0.05)
plot(fit)
```

print.ssc *Print an ssc Object*

Description

Compact one-line summary of a "ssc" estimation result.

Usage

```
## S3 method for class 'ssc'
print(x, ...)
```

Arguments

x	An object of class "ssc", as returned by ssc.
...	Currently unused.

Value

x, invisibly.

ssc *Staggered Synthetic Control Estimation*

Description

Estimate treatment effects in a panel with staggered adoption using the Staggered Synthetic Control (SSC) method of Cao, Lu, and Wu (2020). Returns event-time ATT, overall ATT, heterogeneous treatment effects, and placebo-in-time confidence intervals.

Usage

```
ssc(Y, D, S = NULL, alpha = 0.05)
```

Arguments

Y	Numeric matrix ($N \times T_{total}$) of outcomes. Rows are units, columns are time periods (pre- and post-treatment).
D	Binary matrix ($N \times T_{total}$) of treatment indicators. D[i,t] = 1 if unit i is treated at time t. Treatment must be absorbing (once treated, always treated).
S	Integer or NULL. Number of post-treatment periods to use. If NULL (default), all available post-treatment periods are used.
alpha	Significance level for confidence intervals (default 0.05).

Details

The SSC method proceeds in four steps:

1. **SC weights.** For every unit, estimate synthetic control weights from pre-treatment data.
2. **Treatment structure.** Build the treatment assignment matrices A_s that map heterogeneous effects γ to unit-level outcomes at each post-treatment period.
3. **Estimation.** Solve a GLS-type system to recover $\hat{\gamma}$, then aggregate to event-time or overall ATT via a linear map L.
4. **Inference.** Construct a null distribution by applying the same estimator to rolling windows of pre-treatment residuals (placebo-in-time). Confidence intervals are the $\alpha/2$ and $1-\alpha/2$ quantiles of this distribution, shifted by the point estimate.

Value

An object of class `"ssc"`, a list containing:

att_event Numeric vector of length S: event-time ATT estimates (averaged across units at each event time).

ci_lower_event, ci_upper_event Numeric vectors of length S: lower and upper bounds of $(1-\alpha)$ placebo-in-time confidence intervals. `NA` when $T < S$ (too few pre-treatment periods).

att_overall Scalar: overall ATT (simple average of all heterogeneous effects).

ci_lower_overall, ci_upper_overall Scalar CI bounds for the overall ATT. `NA` when $T < S$.

p_value Two-sided p-value for $H_0 : ATT = 0$ based on the placebo distribution. `NA` when $T < S$.

gamma_hat Numeric vector of length K: heterogeneous treatment effects for every treated (unit, post-period) pair.

te_mat_hat Numeric $N \times S$ matrix of unit-level treatment effects at each post-treatment period.

B_hat Numeric $N \times N$ SC weight matrix.

a_hat Numeric vector of length N: SC intercepts.

u_hat Numeric $N \times T$ matrix of pre-treatment SC residuals.

min_eigenvalue Smallest eigenvalue of the sample analogue of the design matrix $\sum_s A_s' \hat{M} A_s$. Must be positive for the estimator to be well-defined.

index_mat Integer $K \times 3$ matrix. Each row (s, i, e) records the post-treatment period s, unit i, and event time e for one element of $\hat{\gamma}$.

N, T, S, K Panel dimensions.

alpha Significance level used.

References

Cao, J., Lu, C., and Wu, Y. (2020). "Synthetic Control Inference for Staggered Adoption."

Examples

```
set.seed(1)
N <- 5; Ttot <- 15
Y <- matrix(rnorm(N * Ttot), N, Ttot)
D <- matrix(0L, N, Ttot)
D[1, 8:15] <- 1L
D[2, 10:15] <- 1L
result <- ssc(Y, D)
print(result)
summary(result)
```

`ssc_min_eigenvalue` *Compute Smallest Eigenvalue of the SSC Design Matrix*

Description

A diagnostic function that builds the SSC design matrix $\sum_s A_s' \hat{M} A_s$ and returns its smallest eigenvalue. This matrix must be positive definite for SSC estimates to exist.

Usage

```
ssc_min_eigenvalue(Y, D, S = NULL)
```

Arguments

`Y`	Numeric matrix ($N \times T_{total}$) of outcomes.
`D`	Binary matrix ($N \times T_{total}$) of treatment indicators.
`S`	Number of post-treatment periods (or `NULL` for all).

Details

A positive value means the SSC estimator is well-defined; a value near zero warns that identification is weak.

Value

A scalar: the smallest eigenvalue.

Examples

```
set.seed(1)
N <- 10; Ttot <- 8
Y <- matrix(rnorm(N * Ttot), N, Ttot)
D <- matrix(0L, N, Ttot)
D[1:3, 5:Ttot] <- 1L   # units 1-3 treated from period 5
ssc_min_eigenvalue(Y, D, S = 2)
```

summary.ssc *Summarise an ssc Object*

Description

Prints a detailed summary of a `"ssc"` estimation result, including design diagnostics, the overall ATT with confidence interval and p-value, and a table of event-time ATT estimates.

Usage

```
## S3 method for class 'ssc'
summary(object, ...)
```

Arguments

`object`	An object of class `"ssc"`, as returned by `ssc`.
`...`	Currently unused.

Value

`object`, invisibly.

synthetic_control *Synthetic Control Weights for a Single Treated Unit*

Description

Estimate synthetic control weights by solving a constrained quadratic program on demeaned pre-treatment outcomes: minimise $\|\tilde{Y}_1 - \tilde{X}b\|^2$ subject to $\sum b_j = 1,\ b_j \geq 0$, where $\tilde{Y}_1$ and $\tilde{X}$ are time-demeaned series for the treated unit and controls.

Usage

```
synthetic_control(Y)
```

Arguments

`Y`	Numeric matrix ($N \times T$). The first row is the treated unit; remaining rows are donor (control) units. Each column is a pre-treatment time period.

Details

The QP is solved by `solve.QP`. A small ridge term ($10^{-6}I$) is added to the Hessian for numerical stability when T is close to or smaller than $N - 1$.

Value

A list with components

a_hat Scalar intercept $\hat{a} = \bar{Y}_1 - \bar{X}'\hat{b}$.

b_hat Numeric vector of length N. Entry 1 is 0 (the treated unit's self-weight); entries $2, \ldots, N$ are the non-negative weights summing to 1.

Examples

```
set.seed(1)
Y <- matrix(rnorm(5 * 20), 5, 20)  # 5 units, 20 pre-treatment periods
res <- synthetic_control(Y)
res$b_hat   # SC weights for unit 1
```

---

synthetic_control_batch

Synthetic Control Weights for All Units (Batch)

Description

For each unit in turn, treat that unit as the "treated" unit and estimate SC weights from the remaining units. This produces an $N \times N$ weight matrix $\hat{B}$ with zeros on the diagonal.

Usage

```
synthetic_control_batch(Y)
```

Arguments

Y	Numeric matrix ($N \times T$) of pre-treatment outcomes. Rows are units, columns are time periods.

Value

A list with components

a_hat Numeric vector of length N: unit-level intercepts.

B_hat Numeric $N \times N$ matrix of SC weights. Row i contains the weights used to construct the synthetic control for unit i; $B_{ii} = 0$.

Examples

```
set.seed(1)
Y <- matrix(rnorm(5 * 20), 5, 20)
res <- synthetic_control_batch(Y)
res$B_hat   # N x N weight matrix
```

Index