

Package: summclust (via r-universe)

October 10, 2026

Title Module to Compute Influence and Leverage Statistics for
Regression Models with Clustered Errors

Version 0.7.0

Description Module to compute cluster specific information for
regression models with clustered errors, including leverage and
influence statistics. Models of type 'lm' and 'fixest'(from the
'stats' and 'fixest' packages) are supported. 'summclust'
implements similar features as the user-written 'summclust.ado'
Stata module (MacKinnon, Nielsen & Webb, 2022;
<[doi:10.48550/arXiv.2205.03288v1](https://doi.org/10.48550/arXiv.2205.03288v1)>).

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.2.1

Imports utils, dreamerr, MASS, collapse, generics, cli, rlang

Suggests ggplot2, latex2exp, fabricatr, fixest, haven, sandwich,
lmtest, testthat (>= 3.0.0), knitr, rmarkdown, covr

Config/testthat/edition 3

URL <https://s3alfisc.github.io/summclust/>

BugReports <https://github.com/s3alfisc/summclust/issues>

VignetteBuilder knitr

Repository <https://s3alfisc.r-universe.dev>

Date/Publication 2023-06-11 16:55:10 UTC

RemoteUrl <https://github.com/s3alfisc/summclust>

RemoteRef HEAD

RemoteSha 2d25c68eb046028b680fdc8c9827820f7e791827

Contents

plot.summclust	2
summary.summclust	3
summclust	4
summclust.fixest	5
summclust.lm	7
tidy.summclust	9
vcov_CR3J	10
vcov_CR3J.fixest	11
vcov_CR3J.lm	13
Index	15

plot.summclust	<i>Plotting method for objects of type summclust</i>
----------------	--

Description

Plots residual leverage, partial leverage and the leave-one-cluster-out regression coefficients

Usage

```
## S3 method for class 'summclust'
plot(x, ...)
```

Arguments

- x An object of type summclust
- ... other optional function arguments

Details

Note that the function requires ggplot2 to be installed.

Value

- A list containing
 - residual_leverage A ggplot of the residual leverages
 - coef_leverage A ggplot of the coefficient leverages
 - coef_beta A ggplot of the leave-one-out cluster jackknife regression coefficients

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using summclust." arXiv preprint arXiv:2205.03288 (2022).

Examples

```

if(requireNamespace("summclust") && requireNamespace("haven")){

  library(summclust)
  library(haven)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  lm_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

  res <- summclust(
    obj = lm_fit,
    params = c("msp", "union"),
    cluster = ~ind_code,
  )

  plot(res)
}

```

summary.summclust	<i>A summary() method for objects of type summclust</i>
-------------------	---

Description

A summary() method for objects of type summclust

Usage

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

Arguments

object	An object of type summclust
...	misc arguments

Value

The function summary.summclust returns a range of cluster leverage statistics based on an object of type summclust

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using summclust." arXiv preprint arXiv:2205.03288 (2022).

Examples

```
if(requireNamespace("summclust") && requireNamespace("haven")){
  library(summclust)
  library(haven)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  lm_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

  res <- summclust(
    obj = lm_fit,
    params = c("msp", "union"),
    cluster = ~ind_code,
  )

  summary(res)
}
```

summclust

Compute Influence and Leverage Metrics

Description

Compute influence and leverage metrics for clustered inference based on the Cluster Jackknife described in MacKinnon, Nielsen & Webb (2022).

Usage

```
summclust(obj, ...)
```

Arguments

obj	An object of class <code>lm</code> or <code>fixest</code>
...	Other arguments

Value

An object of type `summclust`, including a CRV3 variance-covariance estimate as described in MacKinnon, Nielsen & Webb (2022)

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using `summclust`." arXiv preprint arXiv:2205.03288 (2022).

See Also

[summclust.lm](#), [summclust.fixest](#)

Examples

```
if(requireNamespace("summclust") && requireNamespace("haven")){  
  
  library(summclust)  
  library(haven)  
  
  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")  
  # drop NAs at the moment  
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]  
  nlswork <- na.omit(nlswork)  
  
  lm_fit <- lm(  
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),  
    data = nlswork)  
  
  res <- summclust(  
    obj = lm_fit,  
    params = c("msp", "union"),  
    cluster = ~ind_code,  
  )  
  
  summary(res)  
  tidy(res)  
  plot(res)  
}
```

`summclust.fixest`*Compute Influence and Leverage Metrics for objects of type `fixest`*

Description

Compute influence and leverage metrics for clustered inference based on the Cluster Jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type `fixest`.

Usage

```
## S3 method for class 'fixest'
summclust(
  obj,
  cluster,
  params,
  absorb_cluster_fixef = TRUE,
  type = "CRV3",
  ...
)
```

Arguments

<code>obj</code>	An object of type <code>fixest</code>
<code>cluster</code>	A clustering vector
<code>params</code>	A character vector of variables for which leverage statistics should be computed. If NULL, leverage statistics will be computed for all k model covariates
<code>absorb_cluster_fixef</code>	TRUE by default. Should the cluster fixed effects be projected out? This increases numerical stability and decreases computational costs
<code>type</code>	"CRV3" or "CRV3J" following MacKinnon, Nielsen & Webb
<code>...</code>	other function arguments passed to <code>'vcov'</code>

Value

An object of type `summclust`, including a CRV3 variance-covariance estimate as described in MacKinnon, Nielsen & Webb (2022)

<code>coef_estimates</code>	The coefficient estimates of the linear model.
<code>vcov</code>	A CRV3 or CRV3J variance-covariance matrix estimate as described in MacKinnon, Nielsen & Webb (2022)
<code>leverage_g</code>	A vector of leverages.
<code>leverage_avg</code>	The cluster leverage.
<code>partial_leverage</code>	The partial leverages.
<code>coef_var_leverage_avg</code>	Coefficient of Variation for the leverage statistic
<code>coef_var_leverage_g</code>	Coefficient of Variation for the Partial Leverage Statistics
<code>coef_var_N_G</code>	Coefficient of Variation for the Cluster Sizes.
<code>beta_jack</code>	The jackknifed' leave-on-cluster-out regression coefficients.
<code>params</code>	The input parameter vector <code>'params'</code> .
<code>N_G</code>	The number of clusters-
<code>call</code>	The <code>summclust()</code> function call.
<code>cluster</code>	The names of the clusters.

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using summclust." arXiv preprint arXiv:2205.03288 (2022).

Examples

```
if(requireNamespace("summclust")
  && requireNamespace("haven")
  && requireNamespace("fixest")){

  library(summclust)
  library(haven)
  library(fixest)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  feols_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

  res <- summclust(
    obj = feols_fit,
    params = c("msp", "union"),
    cluster = ~ind_code,
  )

  summary(res)
  tidy(res)
  plot(res)
}
```

summclust.lm

Compute Influence and Leverage Metrics for objects of type lm

Description

Compute influence and leverage metrics for clustered inference based on the Cluster Jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type `lm`.

Usage

```
## S3 method for class 'lm'
summclust(obj, cluster, params, type = "CRV3", ...)
```

Arguments

<code>obj</code>	An object of type <code>lm</code>
<code>cluster</code>	A clustering vector
<code>params</code>	A character vector of variables for which leverage statistics should be computed.
<code>type</code>	"CRV3" or "CRV3J" following MacKinnon, Nielsen & Webb. CRV3 by default
<code>...</code>	other function arguments passed to <code>'vcov'</code>

Value

An object of type `summclust`, including a CRV3 variance-covariance estimate as described in MacKinnon, Nielsen & Webb (2022)

<code>coef_estimates</code>	The coefficient estimates of the linear model.
<code>vcov</code>	A CRV3 or CRV3J variance-covariance matrix estimate as described in MacKinnon, Nielsen & Webb (2022)
<code>leverage_g</code>	A vector of leverages.
<code>leverage_avg</code>	The cluster leverage.
<code>partial_leverage</code>	The partial leverages.
<code>beta_jack</code>	The jackknifed' leave-on-cluster-out regression coefficients.
<code>params</code>	The input parameter vector <code>'params'</code> .
<code>N_G</code>	The number of clusters-
<code>call</code>	The <code>summclust()</code> function call.
<code>cluster</code>	The names of the clusters.

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using `summclust`." *arXiv preprint arXiv:2205.03288* (2022).

Examples

```
if(requireNamespace("summclust") && requireNamespace("haven")){

  library(summclust)
  library(haven)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  lm_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)
```

```

res <- summclust(
  obj = lm_fit,
  cluster = ~ind_code,
  params = c("msp", "union")
)

summary(res)
tidy(res)
plot(res)
}

```

tidy.summclust	<i>S3 method to summarize objects of class boottest into tidy data.frame</i>
----------------	--

Description

Obtain results from a summclust object in a tidy data frame.

Usage

```

## S3 method for class 'summclust'
tidy(x, ...)

```

Arguments

x	An object of class 'summclust'
...	Other arguments

Value

A data.frame containing coefficient estimates, t-statistics, standard errors, p-value, and confidence intervals based on CRV3 variance-covariance matrix and t(G-1) distribution

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using summclust." arXiv preprint arXiv:2205.03288 (2022).

Examples

```

if(requireNamespace("summclust") && requireNamespace("haven")){

  library(summclust)
  library(haven)
}

```

```

nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
# drop NAs at the moment
nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
nlswork <- na.omit(nlswork)

lm_fit <- lm(
  ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
  data = nlswork)

res <- summclust(
  obj = lm_fit,
  params = c("msp", "union"),
  cluster = ~ind_code,
)

tidy(res)
}

```

vcov_CR3J

Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022)

Description

Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022)

Usage

```
vcov_CR3J(obj, ...)
```

Arguments

obj	An object of class <code>lm</code> or <code>fixest</code> computed?
...	misc function argument

Value

An object of type `'vcov_CR3J'`

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using summclust." arXiv preprint arXiv:2205.03288 (2022).

See Also

[vcov_CR3J.lm](#), [vcov_CR3J.fixest](#)

Examples

```
if(requireNamespace("summclust") && requireNamespace("haven")){

  library(summclust)
  library(haven)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  lm_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

  # CRV3 standard errors
  vcov <- vcov_CR3J(
    obj = lm_fit,
    cluster = ~ind_code,
    type = "CRV3"
  )

  # CRV3 standard errors
  vcovJN <- vcov_CR3J(
    obj = lm_fit,
    cluster = ~ind_code,
    type = "CRV3J",
  )
}
```

vcov_CR3J.fixest	<i>Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type fixest</i>
------------------	---

Description

Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type fixest

Usage

```
## S3 method for class 'fixest'
vcov_CR3J(
```

```

    obj,
    cluster,
    type = "CRV3",
    return_all = FALSE,
    absorb_cluster_fixef = TRUE,
    ...
  )

```

Arguments

<code>obj</code>	An object of type <code>fixest</code>
<code>cluster</code>	A clustering vector
<code>type</code>	"CRV3" or "CRV3J" following MacKinnon, Nielsen & Webb. CRV3 by default
<code>return_all</code>	Logical scalar, FALSE by default. Should only the vcov be returned (FALSE) or additional results (TRUE)
<code>absorb_cluster_fixef</code>	TRUE by default. Should the cluster fixed effects be projected out? This increases numerical stability.
<code>...</code>	other function arguments passed to <code>'vcov'</code>

Value

An object of class `vcov_CR3J`

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using `summclust`." arXiv preprint arXiv:2205.03288 (2022).

Examples

```

if(requireNamespace("summclust")
  && requireNamespace("haven")
  && requireNamespace("fixest")){

  library(summclust)
  library(haven)
  library(fixest)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  feols_fit <- feols(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

```

```

# CRV3 standard errors
vcov <- vcov_CR3J(
  obj = feols_fit,
  cluster = ~ind_code,
  type = "CRV3"
)

# CRV3 standard errors
vcovJN <- vcov_CR3J(
  obj = feols_fit,
  cluster = ~ind_code,
  type = "CRV3J",
)
}

```

vcov_CR3J.lm

Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type lm

Description

Compute CRV3 covariance matrices via a cluster jackknife as described in MacKinnon, Nielsen & Webb (2022) for objects of type `lm`

Usage

```

## S3 method for class 'lm'
vcov_CR3J(obj, cluster, type = "CRV3", return_all = FALSE, ...)

```

Arguments

<code>obj</code>	An object of type <code>lm</code>
<code>cluster</code>	A clustering vector
<code>type</code>	"CRV3" or "CRV3J" following MacKinnon, Nielsen & Webb. CRV3 by default
<code>return_all</code>	Logical scalar, FALSE by default. Should only the <code>vcov</code> be returned (FALSE) or additional results (TRUE)
<code>...</code>	other function arguments passed to <code>'vcov'</code>

Value

An object of class `vcov_CR3J`

References

MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb. "Leverage, influence, and the jackknife in clustered regression models: Reliable inference using `summclust`." arXiv preprint arXiv:2205.03288 (2022).

Examples

```
if(requireNamespace("sumclust") && requireNamespace("haven")){

  library(sumclust)
  library(haven)

  nlswork <- read_dta("http://www.stata-press.com/data/r9/nlswork.dta")
  # drop NAs at the moment
  nlswork <- nlswork[, c("ln_wage", "grade", "age", "birth_yr", "union", "race", "msp", "ind_code")]
  nlswork <- na.omit(nlswork)

  lm_fit <- lm(
    ln_wage ~ union + race + msp + as.factor(birth_yr) + as.factor(age) + as.factor(grade),
    data = nlswork)

  # CRV3 standard errors
  vcov <- vcov_CR3J(
    obj = lm_fit,
    cluster = ~ind_code,
    type = "CRV3"
  )

  # CRV3 standard errors
  vcovJN <- vcov_CR3J(
    obj = lm_fit,
    cluster = ~ind_code,
    type = "CRV3J",
  )
}
```

Index

`plot.summclust`, [2](#)

`summary.summclust`, [3](#)

`summclust`, [4](#)

`summclust.fixest`, [5](#), [5](#)

`summclust.lm`, [5](#), [7](#)

`tidy.summclust`, [9](#)

`vcov_CR3J`, [10](#)

`vcov_CR3J.fixest`, [11](#), [11](#)

`vcov_CR3J.lm`, [11](#), [13](#)