Package ‘MDEI’

May 4, 2023

Type Package

Title Implementing the Method of Direct Estimation and Inference

Version 1.0

Date 2023-05-03

Description Causal and statistical inference on an arbitrary treatment
effect curve requires care in both estimation and inference. This
package, implements the Method of Direct Estimation and Inference as introduced in “Estima-
tion and Inference on Nonlinear and Heterogeneous Effects” by Ratkovic and Tingley (2023) <doi:10.1086/723811>. The method takes an outcome, variable of theoretical interest
(treatment), and set of variables and then returns a partial
derivative (marginal effect) of the treatment variable at each point
along with uncertainty intervals. The approach offers two advances.
First, a split-sample approach is used as a guard against over-fitting.
Second, the method uses a data-driven interval derived from conformal
inference, rather than relying on a normality assumption on the error
terms.

License GPL (>= 2)

Depends R (>= 3.6.0)

Imports grDevices, MASS, ranger, Rcpp (>= 1.0.6), splines2

LinkingTo Rcpp, RcppArmadillo

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation yes

Author Marc Ratkovic [aut, cre],
Dustin Tingley [ctb],
Nithin Kavi [aut]

Maintainer Marc Ratkovic <ratkovic@princeton.edu>

Repository CRAN

Date/Publication 2023-05-04 08:30:11 UTC
R topics documented:

- MDEI-package
- coverPlot
- MDEI
- summary.MDEI

Index

MDEI-package

Summary

MDEI estimates effects.

Author(s)

Maintainer: Marc Ratkovic <ratkovic@princeton.edu>
Authors:

- Nithin Kavi <nithinkavi@college.harvard.edu>

Other contributors:

- Dustin Tingley <dtingley@gov.harvard.edu> [contributor]

Description

Coverage Plot for MDEI Object

Usage

coverPlot(
  object,
  xvar = "treat",
  sigval = 0,
  target = "tau",
  colors = c(gray(0.7), gray(0)),
  cex.point = 0.5,
  xlabel = "",
  ylabel = "",
  ...
)
**MDEI**

**Description**

Implements the Method of Direct Estimation and Inference

**Usage**

```r
MDEI(
  y,
  treat,
  X,
  splits = 10,
  alpha = 0.9,
  samplesplit = TRUE,
  conformal = TRUE,
  nthreads.ranger = NULL,
  verbose = TRUE
)
```
Arguments

- **y**: The outcome variable, a vector.
- **treat**: The treatment variable, a vector.
- **X**: A matrix of covariates.
- **splits**: Number of repeated cross-fitting steps to implement.
- **alpha**: The desired level of the confidence band.
- **samplesplit**: Whether to use a sample splitting approach. Default is TRUE.
- **conformal**: Whether to generate a conformal bands or use a critical value from the normal approximation. Default is TRUE.
- **nthreads.ranger**: Number of threads used internally by the ranger function for random forests. Default is NULL.
- **verbose**: An optional logical value. If TRUE information on the number of split samples completed is printed. Default is TRUE.

Value

- **tau.est**: The estimated marginal effect.
- **CIs.tau**: Upper and lower values of conformal confidence band.
- **critical.values**: Conformal critical values.
- **Ey.x**: Mean of outcome given only covariates.
- **coefficients**: The list of all nonparametric bases and the proportion of sample splits that they were selected.
- **internal**: Internal objects used for development and diagnostics.

References


Examples

```r
n <- 100

X <- matrix(rnorm(n*1), nrow = n)
treat <- rnorm(n)
y <- treat^2 + X[,1] + rnorm(n)

# Be sure to run with more splits than this. We recommend
# at least 10-50 initially, for exploratory analyses, with several hundred for
# publication quality. For large sample sizes, these numbers may be adjusted down.
# These are only recommendations.
# Threads are set to 1 to pass CRAN checks, but we suggest leaving it at the default
# which ranger takse as the total number available.
set.seed(1)
m1 <- MDEI(y, treat, X, splits=1, alpha=.9, nthreads.ranger = 1)
```
# Accuracy
\[
\text{cor(m1$tau.est, treat*2)}
\]
\[
\text{cor(m1$theta.est, treat^2)}
\]

# Coverage
\[
\text{mean(apply(m1$CIs.tau-2*treat,1,prod)<0)}
\]

## S3 method for class 'MDEI'

## Arguments

- **object**: An object of class MDEI.
- **features**: Number of spline bases to include.
- **...**: Additional arguments to be passed to plot.

## Value

**coeftable** A table with three columns: the names of selected spline interactions, the average coefficient, and proportion of time it was included in the model. Averages over taken over subsamples in the split sample strategy. Note that the coefficients are interactions between spline interactions that can be accessed through \( \text{obj$internal$Xmat.spline} \).
Index

coverPlot, 2
MDEI, 3
MDEI-package, 2
summary.MDEI, 5