Package 'OTRselect'

May 3, 2025

Type Package

Title Variable Selection for Optimal Treatment Decision

Version 1.3

Date 2025-05-03

Author Wenbin Lu [aut],
Hao Helen Zhang [aut],
Donglin Zeng [aut],
Yuan Geng [aut],
Shannon T. Holloway [aut, cre]

Maintainer Shannon T. Holloway <shannon.t.holloway@gmail.com>

Description A penalized regression framework that can simultaneously estimate the optimal treatment strategy and identify important variables. Appropriate for either censored or uncensored continuous response.

License GPL-2

Depends stats, lars, survival, methods

NeedsCompilation no

Repository CRAN

Date/Publication 2025-05-03 21:40:05 UTC

Contents

`OTRselect-package` *Variable Selection for Optimal Treatment Decision*

Description

A penalized regression framework that can simultaneously estimate the optimal treatment strategy and identify important variables. Appropriate for either censored or uncensored continuous response.

Details

The DESCRIPTION file:

Package:	OTRselect
Type:	Package
Title:	Variable Selection for Optimal Treatment Decision
Version:	1.3
Date:	2025-05-03
Author:	Wenbin Lu [aut], Hao Helen Zhang [aut], Donglin Zeng [aut], Yuan Geng [aut], Shannon T. Holloway [a
Maintainer:	Shannon T. Holloway <shannon.t.holloway@gmail.com>
Authors@R:	c(person(given = "Wenbin", family = "Lu", role = "aut"), person(given = c("Hao", "Helen"), family = "Z
Description:	A penalized regression framework that can simultaneously estimate the optimal treatment strategy and i
License:	GPL-2
Depends:	stats, lars, survival, methods
NeedsCompilation:	no

Index of help topics:

Function `censored` performs variable selection for censored continuous response. Function `uncensored` performs variable selection for uncensored continuous response. Function `Qhat` estimates the restricted mean response given a treatment regime for censored data or the mean response given a treatment regime for uncensored data.

Author(s)

Wenbin Lu [aut], Hao Helen Zhang [aut], Donglin Zeng [aut], Yuan Geng [aut], Shannon T. Holloway [aut, cre]

Maintainer: Shannon T. Holloway <shannon.t.holloway@gmail.com>

References

Lu, W., Zhang, H. H., and Zeng. D. (2013). Variable selection for optimal treatment decision. Statistical Methods in Medical Research, 22, 493–504. PMCID: PMC3303960.

Geng, Y., Lu, W., and Zhang, H. H. (2015). On optimal treatment regimes selection for mean survival time. Statistics in Medicine, 34, 1169–1184. PMCID: PMC4355217.

`censored` *Variable Selection for Optimal Treatment Decision with Censored Survival Times*

Description

A penalized regression framework that can simultaneously estimate the optimal treatment strategy and identify important variables when the response is continuous and censored. This method uses an inverse probability weighted least squares estimation with adaptive LASSO penalty for variable selection.

Usage

```
censored(x, y, a, delta, propen, phi, logY = TRUE,
         intercept = TRUE)
```

Arguments

`x`	Matrix or data.frame of model covariates.
`y`	Vector of response. Note that this data is used to estimate the Kaplan-Meier Curve and should not be log(T).
`a`	Vector of treatment received. Treatments must be coded as integers or numerics that can be recast as integers without loss of information.
`delta`	Event indicator vector. The indicator must be coded as 0/1 where 0=no event and 1=event.
`propen`	Vector or matrix of propensity scores for each treatment. If a vector, the propensity is assumed to be the same for all samples. Column or element order must correspond to the sort order of the treatment variable, i.e., 0,1,2,3,... If the number of columns/elements in propen is one fewer than the total number of treatment options, it is assumed that the base or lowest valued treatment has not been provided.
`phi`	A character {'c' or 'l'} indicating if the constant ('c') or linear ('l') baseline mean function is to be used.
`logY`	TRUE/FALSE indicating if log(y) is to be used for regression.
`intercept`	TRUE/FALSE indicating if an intercept is to be included in phi model.

Value

A list object containing

beta	A vector of the estimated regression coefficients after variable selection.
optTx	The estimated optimal treatment for each sample.

Author(s)

Wenbin Lu, Hao Helen Zhang, Yuan Geng, and Shannon T. Holloway

References

Geng, Y., Lu, W., and Zhang, H. H. (2015). On optimal treatment regimes selection for mean survival time. Statistics in Medicine, 34, 1169–1184. PMCID: PMC4355217.

Examples

```
sigma <- diag(10)
ct <- 0.5^{1L:9L}
rst <- unlist(sapply(1L:9L,function(x){ct[1L:{10L-x}]}))
sigma[lower.tri(sigma)] <- rst
sigma[upper.tri(sigma)] <- t(sigma)[upper.tri(sigma)]

M <- t(chol(sigma))
Z <- matrix(rnorm(1000),10,100)
X <- t(M%*%Z)

A <- rbinom(100,1,0.5)

Y <- rweibull(100,shape=0.5,scale=1)
C <- rweibull(100,shape=0.5,scale=1.5)

delta <- as.integer(C <= Y)

Y[delta > 0.5] <- C[delta>0.5]

dat <- data.frame(X,A,exp(Y),delta)
colnames(dat) <- c(paste("X",1:10,sep=""),"a","y","del")

censored(x = X,
         y = Y,
         a = A,
         delta = delta,
         propen = 0.5,
         phi = "c",
         logY = TRUE,
         intercept = TRUE)
```

Qhat *Mean Response or Restricted Mean Response Given a Treatment Regime*

Description

Estimates the mean response given a treatment regime if data is uncensored. If data is censored, estimates the restricted mean response given a treatment regime.

Usage

```
Qhat(y, a, g, wgt = NULL)
```

Arguments

y	vector of responses. Note if logY = TRUE in censored, this value should also be the logarithm.
a	vector of treatments received.
g	vector of the given treatment regime.
wgt	weights to be used if response is censored.

Value

Returns the estimated mean response or restricted mean response.

Author(s)

Wenbin Lu, Hao Helen Zhang, Donglin Zeng, Yuan Geng, and Shannon T. Holloway

References

Lu, W., Zhang, H. H., and Zeng. D. (2013). Variable selection for optimal treatment decision. Statistical Methods in Medical Research, 22, 493–504. PMCID: PMC3303960.

Geng, Y., Lu, W., and Zhang, H. H. (2015). On optimal treatment regimes selection for mean survival time. Statistics in Medicine, 34, 1169–1184. PMCID: PMC4355217.

Examples

```
y <- rnorm(100)
a <- rbinom(100,1,0.5)
g <- integer(100)

Qhat(y = y, a = a, g = g)
```

uncensored *Variable Selection for Optimal Treatment Decision with Uncensored Continuous Response*

Description

A penalized regression framework that can simultaneously estimate the optimal treatment strategy and identify important variables when the response is continuous and not censored. This method uses an inverse probability weighted least squares estimation with adaptive LASSO penalty for variable selection.

Usage

```
uncensored(x, y, a, propen, phi, intercept = TRUE)
```

Arguments

x	Matrix or data.frame of model covariates.
y	Vector of response. Note that this data is used to estimate the Kaplan-Meier Curve and should not be log(T).
a	Vector of treatment received. Treatments must be coded as integers or numerics that can be recast as integers without loss of information.
propen	Vector or matrix of propensity scores for each treatment. If a vector, the propensity is assumed to be the same for all samples. Column or element order must correspond to the sort order of the treatment variable, i.e., 0,1,2,3,... If the number of columns/elements in propen is one fewer than the total number of treatment options, it is assumed that the base or lowest valued treatment has not been provided.
phi	A character {'c' or 'l'} indicating if the constant ('c') or linear ('l') baseline mean function is to be used.
intercept	TRUE/FALSE indicating if an intercept is to be included in phi model.

Value

A list object containing

beta	A vector of the estimated regression coefficients after variable selection.
optTx	The estimated optimal treatment for each sample.

Author(s)

Wenbin Lu, Hao Helen Zhang, Donglin Zeng, and Shannon T. Holloway

References

Lu, W., Zhang, H. H., and Zeng. D. (2013). Variable selection for optimal treatment decision. Statistical Methods in Medical Research, 22, 493–504. PMCID: PMC3303960.

Examples

```
sigma <- diag(10)
ct <- 0.5^{1L:9L}
rst <- unlist(sapply(1L:9L,function(x){ct[1L:{10L-x}]}))
sigma[lower.tri(sigma)] <- rst
sigma[upper.tri(sigma)] <- t(sigma)[upper.tri(sigma)]

M <- t(chol(sigma))
Z <- matrix(rnorm(1000),10,100)
X <- t(M %*% Z)

gamma1 <- c(1, -1, rep(0,8))
beta <- c(1,1,rep(0,7), -0.9, 0.8)

A <- rbinom(100,1,0.5)

Y <- 1.0 + X %*% gamma1 +
     A*{cbind(1.0,X)%*%beta} + rnorm(100,0,.25)

dat <- data.frame(X,A,Y)

uncensored(x=X,
           y = Y,
           a = A,
           propen = 0.5,
           phi = "c",
           intercept = TRUE)
```

Index