

Package ‘CostTrajectory’

March 2, 2021

Type Package

Title Statistical Modeling of Longitudinal Medical Cost Trajectory

Version 0.1.0

Date 2021-12-02

Author Shikun Wang

Maintainer Shikun Wang <shikun.wang@uth.tmc.edu>

Description Estimating two-dimensional triangular surface of mean medical costs conditional on survival with P-splines.

License GPL (>= 2)

LazyData true

LazyLoad true

RoxygenNote 7.1.1

Encoding UTF-8

Depends MASS, mgcv, Matrix, survival, foreach, plot3D

R topics documented:

CostTrajectory	2
CostTrajectory_checker	3
CostTrajectory_design	4
CostTrajectory_estimate	5
CostTrajectory_PopulationMean	7
CostTrajectory_simulate_data	8
CostTrajectory_survival	8
CostTrajectory_variance	9
plot.CostTrajectory	10
summary.CostTrajectory	11
Index	12

CostTrajectory

Fit longitudinal medical cost trajectory

Description

The CostTrajectory function fits two-dimensional P-splines estimating equations of the input data of degree and order fixed by the user. Specifically tailored to the cost data.

Usage

```
CostTrajectory(
  time,
  surv,
  cost,
  id = NULL,
  status = NULL,
  ndx = 5,
  deg = 2,
  pord = 2,
  lambdas = c(1e-05, 1e-05, 1e-05),
  coefstart = NULL,
  usecensor = TRUE,
  control = list(),
  correlation = "compound",
  varfunc = "spline"
)
```

Arguments

time	vector for the measurement time of costs.
surv	vector for the observed survival time of costs. The length of "surv" should be the same as the number of observations.
cost	vector for the observed monthly costs.
id	vector for the identifier of each subject. The length of "id" should be the same as the number of observations.
status	vector for the survival status at the observed survival time, death (1) or censored (0).
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
pord	vector with the order of differences for each axis. Default: [2, 2].
coefstart	an optional matrix of starting coefficients.
usecensor	a bool to decide whether to use censored cost data, default is TRUE.
control	a list of control parameters. See Details.
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "compound" and "ar1"
varfunc	a character string specifying the variance structure. The following are permitted: "constant", "spline" and "linear"
lambda	vector with smoothing parameters, possibly three for axis (optional).

Value

An object of type 'CostTrajectory'

Examples

```
data <- CostTrajectory_simulate_data(n=50)
#   id time  surv  status  cost
#  1  1    1    1      1    11.03952
#  2  2    1    2      1    12.96982
#  3  2    2    2      1    13.35871
ndx <- 5; deg <- 2; pord <- 2
lambdas <- rep(1e-5,3)
coefstart <- NULL; control <- list()
correlation <- 'compound'; usecensor <- TRUE
CostTrajectory(time=data$time, surv=data$urv, cost=data$Y,
               id=data$id, status=data$status, ndx, deg, pord,
               lambdas, coefstart, usecensor, control, correlation)
```

CostTrajectory_checker

Fit longitudinal medical cost trajectory

Description

The CostTrajectory_checker function checks whether the input for CostTrajectory are valid

Usage

```
CostTrajectory_checker(
  time,
  surv,
  cost,
  id,
  status,
  ndx,
  deg,
  pord,
  lambdas,
  coefstart,
  control,
  correlation,
  varfunc
)
```

Arguments

time	vector for the measurement time of costs.
surv	vector for the observed survival time of costs.
cost	vector for the observed monthly costs.

id	vector for the identifier of each subject. The length of "id" should be the same as the number of observations.
status	vector for the survival status at the observed survival time, death (1) or censored (0).
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
pord	vector with the order of differences for each axis. Default: [2, 2].
lambdas	vector with smoothing parameters, possibly three for axis (optional).
coefstart	an optional matrix of starting coefficients.
control	a list of control parameters. See Details.
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "compound" and "ar1"
varfunc	a character string specifying the variance structure. The following are permitted: "constant", "spline" and "linear"

Value

A list containing CHECKED arguments for the CostTrajectory function

Examples

```
time <- Reduce('c',sapply(1:11, function(x) 1:x))
surv <- id <- Reduce('c',sapply(1:11, function(x) rep(x,x)))
cost <- (time-5)^2*surv/10+rnorm(66,10,1)
status <- c(rep(1,55), rep(0,11))
ndx <- 5; deg <- 2; pord <- 2
lambdas <- rep(1e-5,3)
coefstart <- NULL; control <- list()
correlation <- 'compound'
CostTrajectory_checker(time, surv, cost, id, status, ndx, deg, pord,
                      lambdas, coefstart, control, correlation)
```

CostTrajectory_design *Calculate design matrix to fit medical cost trajectory*

Description

The CostTrajectory_design function generate the design matrix for CostTrajectory function.

Usage

```
CostTrajectory_design(
  time,
  surv,
  cost,
  id,
  status,
  ndx = 5,
  deg = 2,
  usecensor = TRUE,
  surv.out
)
```

Arguments

time	vector for the measurement time of costs.
surv	vector for the observed survival time of costs. The length of "surv" should be the same as the number of observations.
cost	vector for the observed monthly costs.
id	vector for the identifier of each subject. The length of "id" should be the same as the number of observations.
status	vector for the survival status at the observed survival time, death (1) or censored (0).
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
usecensor	a bool to decide whether to use censored cost data, default is TRUE.
pord	vector with the order of differences for each axis. Default: [2, 2].
lambda	vector with smoothing parameters, possibly three for axis (optional).
coefstart	an optional matrix of starting coefficients.
control	a list of control parameters. See Details.
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "compound" and "ar1"

Value

A list containing design arguments for the CostTrajectory function

Examples

```
time <- Reduce('c',sapply(1:11, function(x) 1:x))
surv <- id <- Reduce('c',sapply(1:11, function(x) rep(x,x)))
cost <- (time-5)^2*surv/10+rnorm(66,10,1)
status <- c(rep(1,55), rep(0,11))
MON <- FALSE; TOL1 <- 1E-4; MAX.IT <- 5
ndx <- 5; deg <- 2; pord <- 2
lambda <- 1e-5; usecensor <- FALSE
CostTrajectory_design(time/surv, surv, cost, id, status,
                      ndx, deg, lambda, usecensor, surv.out = 0)
```

CostTrajectory_estimate

Estimate parameters for longitudinal medical cost trajectory

Description

The CostTrajectory_estimate function estimate mean and variance parameters.

Usage

```
CostTrajectory_estimate(
  B,
  y,
  z,
  id,
  status,
  usecensor,
  varfunc,
  theta.old = NULL,
  ndx,
  deg,
  pord,
  lambdas,
  MON,
  MAX.IT,
  correlation,
  LTS = TRUE
)
```

Arguments

B	vector for scaled the measurement time of costs.
y	vector for the scaled observed survival time of costs.
z	vector for the observed monthly costs.
id	vector for the identifier of each subject. The length of "id" should be the same as the number of observations.
status	vector for the survival status at the observed survival time, death (1) or censored (0).
varfunc	a character string specifying the variance structure. The following are permitted: "constant", "spline" and "linear"
theta.old	initial value for model parameter, default is NULL.
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
lambdas	vector with smoothing parameters, possibly three for axis (optional).
MON	a control bool value for monitor status, default is FALSE.
MAX.IT	a value for maximum number of iterations, default is 20
correlation	a character string specifying the correlation structure. The following are permitted: "independence", "compound" and "AR"
LTS	a bool indicator whether to specify long-term survivor group, default is TRUE

Value

A list containing parameter arguments for the CostTrajectory function

Examples

```

surv <- id <- Reduce('c',sapply(1:11, function(x) rep(x,x)))
y <- surv / max(surv)
z <- (time-5)^2*surv/10+rnorm(66,10,1)
status <- c(rep(1,55), rep(0,11))
MON <- FALSE; TOL1 <- 1E-4; MAX.IT <- 5
ndx <- 5; deg <- 2; pord <- 2
varfunc <- 'constant'
correlation <- 'compound'
B <- matrix(rnorm(3696),66,56)
lambdas <- rep(1e-5,3)
CostTrajectory_estimate(B, y, z, id, status, usecensor, varfunc, theta.old = NULL,
                        ndx, deg, lambdas, MON, MAX.IT, correlation, LTS = TRUE)

```

CostTrajectory_PopulationMean

Calculate population mean longitudinal medical cost trajectory

Description

The CostTrajectory_PopulationMean function produces the pointwise averaged estimate based on data from uncensored patients and long-term survivors.

Usage

```
CostTrajectory_PopulationMean(dat)
```

Arguments

dat a data frame storing cost and survival data.

Value

A data frame of estimated cost trajectory

Examples

```

data <- CostTrajectory_simulate_data(n=50)
CostTrajectory_PopulationMean(data)

```

CostTrajectory_simulate_data

Simulate longitudinal medical cost trajectory

Description

The CostTrajectory_simulate_data function simulate longitudinal medical cost data w/o error to train / test the model

Usage

```
CostTrajectory_simulate_data(n = 1000, seed = 123, MON = F, test = F)
```

Arguments

n	number of subjects. maximum follow-up month if for testing (test=TRUE).
seed	seed of random error generator for training (test=FALSE).
MON	a control bool value for monitor status, default is FALSE.
test	a bool value to indicate whether to simulate train (=FALSE) / test (=TRUE) data. Train data has multivariate normal error and weibull survival. Test data is observed by month without error.

Value

A data frame containing valid input for CostTrajectory function.

Examples

```
CostTrajectory_simulate_data(n=10, test=T)
```

CostTrajectory_survival

Calculate survival function to fit medical cost trajectory

Description

The CostTrajectory_survival function estimates the hazard function and the corresponding weight matrix required to calculate the design matrix in CostTrajectory function.

Usage

```
CostTrajectory_survival(
  surv0,
  status0,
  MON,
  TOL1,
  MAX.IT,
  ndx,
  deg,
```



```

    lambda = NULL,
    usecensor = TRUE
  )

```

Arguments

surv0	vector for the observed survival time of costs.
status0	vector for the survival status at the observed survival time, death (1) or censored (0).
MON	monitor indicator, default is FALSE.
TOL1	tolerance for convergence criteria, default is 1E-4.
MAX.IT	maximum number of iterations, default is 50.
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
usecensor	a bool to decide whether to use censored cost data, default is TRUE.

Value

A list containing survival arguments for the CostTrajectory function

Examples

```

surv0 <- 1:11
status0 <- c(rep(1,10),1)
MON <- FALSE; TOL1 <- 1E-4; MAX.IT <- 5
ndx <- 5; deg <- 2; pord <- 2
lambda <- 1e-5; usecensor <- TRUE
CostTrajectory_survival(surv0, status0, MON, TOL1, MAX.IT, ndx, deg, lambda, usecensor)

```

CostTrajectory_variance

Calculate survival function to fit medical cost trajectory

Description

The CostTrajectory_variance function calculate the variance function for CostTrajectory function.

Usage

```

CostTrajectory_variance(
  B,
  id,
  z,
  status,
  y,
  dB,
  surv.out,
  varfunc,
  theta,

```

```

    lambdas,
    correlation,
    ndx,
    deg
  )

```

Arguments

id	vector for the identifier of each subject. The length of "id" should be the same as the number of observations.
ndx	vector with the number of internal knots-1 for each axis, default is [5, 5].
deg	vector with the degree of the B-splines for each axis default is [2, 2].
surv0	vector for the observed survival time of costs.
status0	vector for the survival status at the observed survival time, death (1) or censored (0).
MON	monitor indicator, default is FALSE.
TOL1	tolerance for convergence criteria, default is 1E-4.
usecensor	a bool to decide whether to use censored cost data, default is TRUE.
MAX.IT	maximum number of iterations, default is 50.

Value

A list containing survival arguments for the CostTrajectory function

Examples

```

surv0 <- 1:11
status0 <- c(rep(1,10),1)
MON <- FALSE; TOL1 <- 1E-4; MAX.IT <- 5
ndx <- 5; deg <- 2; pord <- 2
lambda <- 1e-5; usecensor <- TRUE
CostTrajectory_variance(surv0, status0, MON, TOL1, MAX.IT, ndx, deg, lambda, usecensor)

```

plot.CostTrajectory *Plot longitudinal medical cost trajectory*

Description

The plot.CostTrajectory function produces three plots for the CostTrajectory object.

Usage

```

## S3 method for class 'CostTrajectory'
plot(object, ...)

```

Arguments

object an object of class "CostTrajectory" from function CostTrajectory.

`summary.CostTrajectory`*Summarize longitudinal medical cost trajectory*

Description

The `print.summary.CostTrajectory` function produces summary statistics for the `CostTrajectory` object.

Usage

```
## S3 method for class 'CostTrajectory'  
summary(object, digits = max(3, getOption("digits") - 3), ...)
```

Arguments

`object` an object of class "CostTrajectory" from function `CostTrajectory`.

Index

`CostTrajectory`, [2](#)
`CostTrajectory_checker`, [3](#)
`CostTrajectory_design`, [4](#)
`CostTrajectory_estimate`, [5](#)
`CostTrajectory_PopulationMean`, [7](#)
`CostTrajectory_simulate_data`, [8](#)
`CostTrajectory_survival`, [8](#)
`CostTrajectory_variance`, [9](#)

`plot.CostTrajectory`, [10](#)

`summary.CostTrajectory`, [11](#)