Package 'psbcGroup'

July 2, 2014

Type Package

Title Penalized semi-parametric Bayesian Cox (PSBC) models with shrinkage and grouping priors

Version 1.0

Date 2013-04-05

Author Kyu Ha Lee, Sounak Chakraborty, (Tony) Jianguo Sun

Maintainer Kyu Ha Lee <klee@hsph.harvard.edu>

Description Algorithms for fitting the PSBC models with shrinkage and grouping priors

License GPL (>= 2)

Depends bayesm, LearnBayes, SuppDists, mvtnorm, R (>= 2.12.2)

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2013-04-05 21:00:23

R topics documented:

psbcEN *The function to fit the PSBC model with elastic net prior*

Description

The function to fit the PSBC model with elastic net prior

Usage

```
psbcEN(survObj, priorPara, initial, rw=FALSE, mcmcPara, num.reps,
thin, chain = 1, save = 1000)
```

Arguments

survObj	The list containing observed data from n subjects; t, di, x
priorPara	The list containing prior parameter values; eta0, kappa0, c0, r1, r2, delta1, delta2, s
initial	The list containing the starting values of the parameters; beta.ini, lambda1Sq, lambda2, sigmaSq, tauSq, h
rw	When setting to "TRUE", the conventional random walk Metropolis Hastings algorithm is used. Otherwise, the mean and the variance of the proposal density is updated using the jumping rule described in Lee et al. (2013).
mcmcPara	The list containing the values of options for Metropolis-Hastings step for β; numBeta, beta.prop.var
num.reps	the number of iterations of the chain
thin	thinning
chain	the numeric name of chain in the case when running multiple chains.
save	frequency of storing the results in .Rdata file. For example, by setting "save = 1000", the algorithm saves the results every 1000 iterations.

Details

t	a vector of n times to the event
di	a vector of n censoring indicators for the event time (1=event occurred, 0=censored)
x	covariate matrix, n observations by p variables
eta0	scale parameter of gamma process prior for the cumulative baseline hazard, $eta0 > 0$
kappa0	shape parameter of gamma process prior for the cumulative baseline hazard, $kappa0 > 0$
c0	the confidence parameter of gamma process prior for the cumulative baseline hazard, $c0 > 0$
r1	the shape parameter of the gamma prior for λ_1^2
r2	the shape parameter of the gamma prior for λ_2
delta1	the rate parameter of the gamma prior for λ_1^2
delta2	the rate parameter of the gamma prior for λ_2
s	the set of time partitions for specification of the cumulative baseline hazard function
beta.ini	the starting values for β

lambda1Sq	the starting value for λ_1^2
lambda2	the starting value for λ_2
sigmaSq	the starting value for σ^2
tauSq	the starting values for τ^2
h	the starting values for h
numBeta	the number of components in β to be updated at one iteration
beta.prop.var	the variance of the proposal density for β when rw is set to "TRUE"

Value

psbcEN returns an object of class psbcEN

beta.p	posterior samples for β
h.p	posterior samples for h
tauSq.p	posterior samples for τ^2
mcmcOutcome	The list containing posterior samples for the remaining model parameters

Note

If the prespecified value of save is less than that of num.reps, the results are saved as .Rdata file under the directory working directory/mcmcOutcome.

Author(s)

Kyu Ha Lee, Sounak Chakraborty, (Tony) Jianguo Sun

References

Lee, K. H., Chakraborty, S., and Sun, J. (2011). Bayesian Variable Selection in Semiparametric Proportional Hazards Model for High Dimensional Survival Data. *The International Journal of Biostatistics*, 7, 21.
Lee, K. H., Chakraborty, S., and Sun, J. (2013). Survival Prediction and Variable Selection with Simultaneous Shrinkage and Grouping Priors, available from the authors upon request.

Examples

```
# generate some survival data

set.seed(204542)

p = 20
n = 100
beta.true <- c(rep(4, 10), rep(0, (p-10)))
```

```
CovX<- diag(0.1, p)

survObj  <- list()
survObj$x <- apply(rmvnorm(n, sigma=CovX, method="chol"), 2, scale)

pred <- as.vector(exp(rowSums(scale(survObj$x, center = FALSE, scale = 1/beta.true))))

t  <- rexp(n, rate = pred)
cen <- runif(n, 0, 8)
survObj$t  <- pmin(t, cen)
survObj$di  <- as.numeric(t <= cen)

priorPara  <- list()
priorPara$eta0  <- 1
priorPara$kappa0  <- 1
priorPara$c0  <- 2
priorPara$r1 <- 0.1
priorPara$r2 <- 1
priorPara$delta1 <- 0.1
priorPara$delta2 <- 1
priorPara$s <- sort(survObj$t[survObj$di == 1])
priorPara$s <- c(priorPara$s, 2 * max(survObj$t) -
max(survObj$t[-which(survObj$t==max(survObj$t))]))
priorPara$J <- length(priorPara$s)

mcmcPara <- list()
mcmcPara$numBeta <- p
mcmcPara$beta.prop.var <- 1

initial <- list()
initial$beta.ini <- rep(0.5, p)
initial$lambda1Sq <- 1
initial$lambda2 <- 1
initial$sigmaSq <- runif(1, 0.1, 10)
initial$tauSq <- rexp(p, rate = initial$lambda1Sq/2)
initial$h <- rgamma(priorPara$J, 1, 1)

rw = FALSE
num.reps = 5
chain = 1
thin = 5
save = 5

fitEN <- psbcEN(survObj, priorPara, initial, rw=FALSE, mcmcPara,
num.reps, thin, chain, save)
```

---

psbcFL *The function to fit the PSBC model with fused lasso prior*

Description

The function to fit the PSBC model with fused lasso prior

Usage

```
psbcFL(survObj, priorPara, initial, rw=FALSE, mcmcPara, num.reps,
thin, chain = 1, save = 1000)
```

Arguments

survObj	The list containing observed data from n subjects; t, di, x
priorPara	The list containing prior parameter values; eta0, kappa0, c0, r1, r2, delta1, delta2, s
initial	The list containing the starting values of the parameters; beta.ini, lambda1Sq, lambda2Sq, sigmaSq, tauSq, h, wSq
rw	When setting to "TRUE", the conventional random walk Metropolis Hastings algorithm is used. Otherwise, the mean and the variance of the proposal density is updated using the jumping rule described in Lee et al. (2013).
mcmcPara	The list containing the values of options for Metropolis-Hastings step for β; numBeta, beta.prop.var
num.reps	the number of iterations of the chain
thin	thinning
chain	the numeric name of chain in the case when running multiple chains.
save	frequency of storing the results in .Rdata file. For example, by setting "save = 1000", the algorithm saves the results every 1000 iterations.

Details

t	a vector of n times to the event
di	a vector of n censoring indicators for the event time (1=event occurred, 0=censored)
x	covariate matrix, n observations by p variables
eta0	scale parameter of gamma process prior for the cumulative baseline hazard, $eta0 > 0$
kappa0	shape parameter of gamma process prior for the cumulative baseline hazard, $kappa0 > 0$
c0	the confidence parameter of gamma process prior for the cumulative baseline hazard, $c0 > 0$
r1	the shape parameter of the gamma prior for λ_1^2
r2	the shape parameter of the gamma prior for λ_2^2
delta1	the rate parameter of the gamma prior for λ_1^2
delta2	the rate parameter of the gamma prior for λ_2^2
s	the set of time partitions for specification of the cumulative baseline hazard function
beta.ini	the starting values for β
lambda1Sq	the starting value for λ_1^2
lambda2Sq	the starting value for λ_2^2
sigmaSq	the starting value for σ^2
tauSq	the starting values for τ^2
h	the starting values for h
wSq	the starting values for w^2

numBeta	the number of components in β to be updated at one iteration
beta.prop.var	the variance of the proposal density for β when rw is set to "TRUE"

Value

psbcFL returns an object of class psbcFL

beta.p	posterior samples for β
h.p	posterior samples for h
tauSq.p	posterior samples for τ^2
mcmcOutcome	The list containing posterior samples for the remaining model parameters

Note

If the prespecified value of save is less than that of num.reps, the results are saved as .Rdata file under the directory working directory/mcmcOutcome.

Author(s)

Kyu Ha Lee, Sounak Chakraborty, (Tony) Jianguo Sun

References

Lee, K. H., Chakraborty, S., and Sun, J. (2011). Bayesian Variable Selection in Semiparametric Proportional Hazards Model for High Dimensional Survival Data. *The International Journal of Biostatistics*, 7, 21.
Lee, K. H., Chakraborty, S., and Sun, J. (2013). Survival Prediction and Variable Selection with Simultaneous Shrinkage and Grouping Priors, available from the authors upon request.

Examples

```
# generate some survival data

set.seed(204542)

p = 20
n = 100
beta.true <- c(rep(4, 10), rep(0, (p-10)))

CovX<- diag(0.1, p)

survObj  <- list()
survObj$x <- apply(rmvnorm(n, sigma=CovX, method="chol"), 2, scale)

pred <- as.vector(exp(rowSums(scale(survObj$x, center = FALSE, scale = 1/beta.true))))
```

```
t  <- rexp(n, rate = pred)
cen <- runif(n, 0, 8)
survObj$t  <- pmin(t, cen)
survObj$di  <- as.numeric(t <= cen)

priorPara  <- list()
priorPara$eta0  <- 2
priorPara$kappa0  <- 2
priorPara$c0  <- 2
priorPara$r1 <- 0.5
priorPara$r2 <- 0.5
priorPara$delta1 <- 0.0001
priorPara$delta2 <- 0.0001
priorPara$s <- sort(survObj$t[survObj$di == 1])
priorPara$s <- c(priorPara$s, 2 * max(survObj$t) -
max(survObj$t[-which(survObj$t==max(survObj$t))]))
priorPara$J <- length(priorPara$s)

mcmcPara <- list()
mcmcPara$numBeta <- p
mcmcPara$beta.prop.var <- 1

initial <- list()
initial$beta.ini <- rep(0.5, p)
initial$lambda1Sq <- 1
initial$lambda2Sq <- 1
initial$sigmaSq <- runif(1, 0.1, 10)
initial$tauSq <- rexp(p, rate = initial$lambda1Sq/2)
initial$h <- rgamma(priorPara$J, 1, 1)
initial$wSq   <- rexp((p-1), rate = initial$lambda2Sq/2)

rw = FALSE
num.reps = 5
chain = 1
thin = 5
save = 5

fitFL <- psbcFL(survObj, priorPara, initial, rw=FALSE, mcmcPara,
num.reps, thin, chain, save)
```

---

psbcGL *The function to fit the PSBC model with group lasso prior*

Description

The function to fit the PSBC models with group lasso and Bayesian lasso priors

Usage

```
psbcGL(survObj, priorPara, initial, rw=FALSE, mcmcPara, num.reps,
thin, chain = 1, save = 1000)
```

Arguments

survObj	The list containing observed data from n subjects; t, di, x
priorPara	The list containing prior parameter values; eta0, kappa0, c0, r, delta, s, groupInd
initial	The list containing the starting values of the parameters; beta.ini, lambdaSq, sigmaSq, tauSq, h
rw	When setting to "TRUE", the conventional random walk Metropolis Hastings algorithm is used. Otherwise, the mean and the variance of the proposal density is updated using the jumping rule described in Lee et al. (2013).
mcmcPara	The list containing the values of options for Metropolis-Hastings step for β; numBeta, beta.prop.var
num.reps	the number of iterations of the chain
thin	thinning
chain	the numeric name of chain in the case when running multiple chains.
save	frequency of storing the results in .Rdata file. For example, by setting "save = 1000", the algorithm saves the results every 1000 iterations.

Details

t	a vector of n times to the event
di	a vector of n censoring indicators for the event time (1=event occurred, 0=censored)
x	covariate matrix, n observations by p variables
eta0	scale parameter of gamma process prior for the cumulative baseline hazard, $eta0 > 0$
kappa0	shape parameter of gamma process prior for the cumulative baseline hazard, $kappa0 > 0$
c0	the confidence parameter of gamma process prior for the cumulative baseline hazard, $c0 > 0$
r	the shape parameter of the gamma prior for λ^2
delta	the rate parameter of the gamma prior for λ^2
s	the set of time partitions for specification of the cumulative baseline hazard function
groupInd	a vector of p group indicator for each variable
beta.ini	the starting values for β
lambdaSq	the starting value for λ^2
sigmaSq	the starting value for σ^2
tauSq	the starting values for τ^2
h	the starting values for h
numBeta	the number of components in β to be updated at one iteration
beta.prop.var	the variance of the proposal density for β when rw is set to "TRUE"

Value

psbcGL returns an object of class psbcGL

beta.p	posterior samples for β
h.p	posterior samples for h
tauSq.p	posterior samples for τ^2
mcmcOutcome	The list containing posterior samples for the remaining model parameters

Note

To fit the PSBC model with the ordinary Bayesian lasso prior (Lee et al., 2011), groupInd needs to be set to 1:p. If the prespecified value of save is less than that of num.reps, the results are saved as .Rdata file under the directory working directory/mcmcOutcome.

Author(s)

Kyu Ha Lee, Sounak Chakraborty, (Tony) Jianguo Sun

References

Lee, K. H., Chakraborty, S., and Sun, J. (2011). Bayesian Variable Selection in Semiparametric Proportional Hazards Model for High Dimensional Survival Data. *The International Journal of Biostatistics*, 7, 21.
Lee, K. H., Chakraborty, S., and Sun, J. (2013). Survival Prediction and Variable Selection with Simultaneous Shrinkage and Grouping Priors, available from the authors upon request.

Examples

```
# generate some survival data

set.seed(204542)

p = 20
n = 100
beta.true <- c(rep(4, 10), rep(0, (p-10)))

CovX<-matrix(0,p,p)

for(i in 1:10){
for(j in 1:10){
CovX[i,j] <- 0.5^abs(i-j)
}
}

diag(CovX) <- 1

survObj  <- list()
survObj$x <- apply(rmvnorm(n, sigma=CovX, method="chol"), 2, scale)
```

```
pred <- as.vector(exp(rowSums(scale(survObj$x, center = FALSE, scale = 1/beta.true))))

t  <- rexp(n, rate = pred)
cen <- runif(n, 0, 8)
survObj$t  <- pmin(t, cen)
survObj$di  <- as.numeric(t <= cen)

priorPara  <- list()
priorPara$eta0  <- 1
priorPara$kappa0  <- 1
priorPara$c0  <- 2
priorPara$r <- 0.5
priorPara$delta <- 0.0001
priorPara$s <- sort(survObj$t[survObj$di == 1])
priorPara$s <- c(priorPara$s, 2 * max(survObj$t) -
max(survObj$t[-which(survObj$t==max(survObj$t))]))
priorPara$J <- length(priorPara$s)
priorPara$groupInd <- c(rep(1,10),2:11)

mcmcPara <- list()
mcmcPara$numBeta <- p
mcmcPara$beta.prop.var <- 1

initial <- list()
initial$beta.ini <- rep(0.5, p)
initial$lambdaSq <- 1
initial$sigmaSq <- runif(1, 0.1, 10)
initial$tauSq <- rexp(length(unique(priorPara$groupInd)), rate = initial$lambdaSq/2)
initial$h <- rgamma(priorPara$J, 1, 1)

rw = FALSE
num.reps = 5
chain = 1
thin = 5
save = 5

fitGL <- psbcGL(survObj, priorPara, initial, rw=FALSE, mcmcPara,
num.reps, thin, chain, save)
```

---

psbcGroup *Penalized semi-parametric Bayesian Cox (PSBC) models with shrinkage and grouping priors*

Description

The package provides algorithms for fitting PSBC models with elastic net, fused lasso, and group lasso priors

Details

psbcEN	The function to fit the PSBC model with elastic net prior
psbcFL	The function to fit the PSBC model with fused lasso prior
psbcGL	The function to fit the PSBC models with group lasso and Bayesian lasso priors

Package:	psbcGroup
Type:	Package
Version:	1.0
Date:	2013-04-05
License:	GPL (>= 2)
LazyLoad:	yes

Author(s)

Kyu Ha Lee, Sounak Chakraborty, (Tony) Jianguo Sun
Maintainer: Kyu Ha Lee <klee@hsph.harvard.edu>

References

Lee, K. H., Chakraborty, S., and Sun, J. (2011). Bayesian Variable Selection in Semiparametric Proportional Hazards Model for High Dimensional Survival Data. *The International Journal of Biostatistics*, 7, 21.

Lee, K. H., Chakraborty, S., and Sun, J. (2013). Survival Prediction and Variable Selection with Simultaneous Shrinkage and Grouping Priors, available from the authors upon request.

Index