

Package ‘BeSS’

April 22, 2021

Type Package

Title Best Subset Selection in Linear, Logistic and CoxPH Models

Version 2.0.3

Date 2021-04-21

Author Canhong Wen [aut, cre],
Aijun Zhang [aut],
Shijie Quan [aut],
Xueqin Wang [aut]

Maintainer Canhong Wen <wencanhong@gmail.com>

Description An implementation of best subset selection in generalized linear model and Cox proportional hazard model via the primal dual active set algorithm proposed by Wen, C., Zhang, A., Quan, S. and Wang, X. (2020) <doi:10.18637/jss.v094.i04>. The algorithm formulates coefficient parameters and residuals as primal and dual variables and utilizes efficient active set selection strategies based on the complementarity of the primal and dual variables.

License GPL-3

Depends R (>= 3.0.0)

Imports Rcpp(>= 0.12.6), Matrix(>= 1.2-6), glmnet, survival

LinkingTo Rcpp, RcppEigen

NeedsCompilation yes

Repository CRAN

Date/Publication 2021-04-22 05:50:02 UTC

R topics documented:

aic	2
bess	3
bess.one	7
coef.bess	10
coef.bess.one	11
deviance.bess	12

deviance.bess.one	13
gen.data	14
gravier	16
logLik.bess	16
logLik.bess.one	17
plot.bess	18
predict.bess	19
predict.bess.one	20
print.bess	21
print.bess.one	22
prostate	23
SAheart	24
summary.bess	24
summary.bess.one	25
Index	27

aic	<i>Extract the IC from a "bess" object.</i>
-----	---

Description

These functions are used by `bess` to compute Information Criteria from a fitted model object.

Usage

```
aic(object,...)
bic(object,...)
ebic(object,...)
```

Arguments

<code>object</code>	Output from the <code>bess</code> function or the <code>bess.one</code> function.
<code>...</code>	Additional arguments affecting the predictions produced.

Value

The value of Information Criteria extracted from the "bess" object.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
aic(fit)
bic(fit)
ebic(fit)
```

bess	<i>Best subset selection</i>
------	------------------------------

Description

Best subset selection for generalized linear model and Cox's proportional model.

Usage

```
bess(x, y, family = c("gaussian", "binomial", "cox"),
     method = "gsection", s.min = 1,
     s.max,
     s.list,
     K.max = 20,
     max.steps = 15,
     glm.max = 1e6,
     cox.max = 20,
     factor = NULL,
     epsilon = 1e-4,
     weights=rep(1,nrow(x)))
```

Arguments

x	Input matrix, of dimension $n \times p$; each row is an observation vector.
y	Response variable, of length n . For family="binomial" should be a factor with two levels. For family="cox", y should be a two-column matrix with columns named 'time' and 'status'.
family	One of the GLM or Cox models. Either "gaussian", "binomial", or "cox", depending on the response.
method	Methods to be used to select the optimal model size. For method = "sequential", we solve the best subset selection problem for each s in $1, 2, \dots, s_{max}$. At each model size s , we run the bess function with a warm start from the last solution with model size $s - 1$. For method = "gsection", we solve the best subset selection problem with a range non-continuous model sizes.

s.min	The minimum value of model sizes. Only used for method = "gsection". Default is 1.
s.max	The maximum value of model sizes. Only used for method = "gsection". Default is $\min p, n/\log(n)$.
s.list	A list of sequential value representing the model sizes. Only used for method = "sequential". Default is $(1, \min p, n/\log(n))$.
K.max	The maximum iterations used for method = "gsection"
max.steps	The maximum number of iterations in bess function. In linear regression, only a few steps can gurantee the convergence. Default is 15.
glm.max	The maximum number of iterations for solving the maximum likelihood problem on the active set at each step in the primal dual active set algorithm. Only used in the logistic regression for family="binomial". Default is 1e6.
cox.max	The maximum number of iterations for solving the maximum partial likelihood problem on the active set at each step in the primal dual active set algorithm. Only used in Cox's model for family="cox". Default is 20.
factor	Which variable to be factored. Should be NULL or a numeric vector.
epsilon	The tolerance for an early stoping rule in the method "sequential". The early stopping rule is defined as $\ Y - X\beta\ /n \leq \epsilon$.
weights	Observation weights. Default is 1 for each observation

Details

The best subset selection problem with model size s is

$$\min_{\beta} -2\log L(\beta) \quad \text{s.t.} \quad \|\beta\|_0 \leq s.$$

In the GLM case, $\log L(\beta)$ is the log-likelihood function; In the Cox model, $\log L(\beta)$ is the log parital likelihood function.

For each candiate model size, the best subset selection problem is solved by the primal dual active set (PDAS) algorithm, see Wen et al(2017) for details. This algorithm utilizes an active set updating strategy via primal and dual vairables and fits the sub-model by exploiting the fact that their support set are non-overlap and complementary. For the case of method = "sequential", we run the PDAS algorithm for a list of sequential model sizes and use the estimate from last iteration as a warm start. For the case of method = "gsection", a golden section search technique is adopted to efficiently determine the optimal model size.

Value

A list with class attribute 'bess' and named components:

family	Types of the model: "bess_gaussian" for linear model, "bess_binomial" for logistic model and "bess_cox" for Cox model.
beta	The best fitting coefficients of size $s = 0, 1, \dots, p$ with the smallest loss function.
lambda	The lambda value in the Lagrangian form of the best subset selection problem with model size of s .

bestmodel	The best fitted model, the class of which is "lm", "glm" or "coxph"
deviance	The value of $-2 \times \log L$.
nulldeviance	The value of $-2 \times \log L$ for null model.
AIC	The value of $-2 \times \log L + 2\ \beta\ _0$.
BIC	The value of $-2 \times \log L + \log(n)\ \beta\ _0$.
EBIC	The value of $-2 \times \log L + (\log(n) + 2 \times \log(p))\ \beta\ _0$.
factor	Which variable to be factored. Should be NULL or a numeric vector.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess.one](#), [plot.bess](#), [predict.bess](#).

Examples

```
#-----linear model-----#
# Generate simulated data
n <- 500
p <- 20
K <-10
sigma <- 1
rho <- 0.2
data <- gen.data(n, p, family = "gaussian", K, rho, sigma)

# Best subset selection
fit1 <- bess(data$x, data$y, family = "gaussian")
print(fit1)
#coef(fit1, sparse=TRUE) # The estimated coefficients
bestmodel <- fit1$bestmodel
#summary(bestmodel)

# Plot solution path and the loss function
plot(fit1, type = "both", breaks = TRUE)

## Not run:
#-----logistic model-----#

# Generate simulated data
data <- gen.data(n, p, family="binomial", 5, rho, sigma)

# Best subset selection
```

```

fit2 <- bess(data$x, data$y, s.list = 1:10, method = "sequential",
             family = "binomial", epsilon = 0)
print(fit2)
#coef(fit2, sparse = TRUE)
bestmodel <- fit2$bestmodel
#summary(bestmodel)

# Plot solution path and the loss function
plot(fit2, type = "both", breaks = TRUE, K = 5)

#-----cox model-----#

# Generate simulated data
data <- gen.data(n, p, 5, rho, sigma, c = 10, family = "cox", scal = 10)

# Best subset selection
fit3 <- bess(data$x, data$y, s.list = 1:10, method = "sequential",
             family = "cox")
print(fit3)
#coef(fit3, sparse = TRUE)
bestmodel <- fit3$bestmodel
#summary(bestmodel)

# Plot solution path and the loss function
plot(fit3, type = "both", breaks = TRUE, K = 5)

#-----High dimensional linear models-----#

p <- 1000
data <- gen.data(n, p, family = "gaussian", K, rho, sigma)

# Best subset selection
fit <- bess(data$x, data$y, method="sequential", family = "gaussian", epsilon = 1e-12)

# Plot solution path
plot(fit, type = "both", breaks = TRUE, K = 10)

data("prostate")
x = prostate[,-9]
y = prostate[,9]

fit.group = bess(x, y, s.list = 1:ncol(x), factor = c("gleason"))

#-----SAheart-----#
data("SAheart")
y = SAheart[,5]
x = SAheart[,-5]
x$ldl[x$ldl<5] = 1
x$ldl[x$ldl>=5&x$ldl<10] = 2
x$ldl[x$ldl>=10] = 3

```

```
fit.group = bess(x, y, s.list = 1:ncol(x), factor = c("ld1"), family = "binomial")
## End(Not run)
```

bess.one

Best subset selection with a specified model size

Description

Best subset selection with a specified model size for generalized linear models and Cox's proportional hazard model.

Usage

```
bess.one(x, y, family = c("gaussian", "binomial", "cox"),
         s = 1,
         max.steps = 15,
         glm.max = 1e6,
         cox.max = 20,
         factor = NULL,
         weights = rep(1,nrow(x)),
         normalize = TRUE)
```

Arguments

x	Input matrix, of dimension $n \times p$; each row is an observation vector.
y	Response variable, of length n . For family = "gaussian", y should be a vector with continuous values. For family = "binomial", y should be a factor with two levels. For family = "cox", y should be a two-column matrix with columns named 'time' and 'status'.
s	Size of the selected model. It controls number of nonzero coefficients to be allowed in the model.
family	One of the distribution function for GLM or Cox models. Either "gaussian", "binomial", or "cox", depending on the response.
max.steps	The maximum number of iterations in the primal dual active set algorithm. In most cases, only a few steps can guarantee the convergence. Default is 15.
glm.max	The maximum number of iterations for solving the maximum likelihood problem on the active set. It occurs at each step in the primal dual active set algorithm. Only used in the logistic regression for family = "binomial". Default is $1e + 6$.
cox.max	The maximum number of iterations for solving the maximum partial likelihood problem on the active set. It occurs at each step in the primal dual active set algorithm. Only used in Cox model for family = "cox". Default is 20.
weights	Observation weights. Default is 1 for each observation
factor	Which variable to be factored. Should be NULL or a numeric vector.
normalize	Whether to normalize x or not. Default is TRUE.

Details

Given a model size s , we consider the following best subset selection problem:

$$\min_{\beta} -2\log L(\beta); s.t. \|\beta\|_0 = s.$$

In the GLM case, $\log L(\beta)$ is the log-likelihood function; In the Cox model, $\log L(\beta)$ is the log partial likelihood function.

The best subset selection problem is solved by the primal dual active set algorithm, see Wen et al. (2017) for details. This algorithm utilizes an active set updating strategy via primal and dual variables and fits the sub-model by exploiting the fact that their support set are non-overlap and complementary.

Value

A list with class attribute 'bess.one' and named components:

type	Types of the model: "bess_gaussian" for linear model, "bess_binomial" for logistic model and "bess_cox" for Cox model
beta	The best fitting coefficients with the smallest loss function given the model size s .
lambda	The estimated lambda value in the Lagrangian form of the best subset selection problem with model size s .
bestmodel	The best fitted model, the class of which is "lm", "glm" or "coxph"
deviance	The value of $-2 * \log L(\beta)$.
nulldeviance	The value of $-2 * \log L(\beta)$ for null model.
factor	Which variable to be factored. Should be NULL or a numeric vector.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [plot.bess](#), [predict.bess](#).

Examples

```
#-----linear model-----#
# Generate simulated data

n <- 500
```



```

p <- 20
K <-10
sigma <- 1
rho <- 0.2
data <- gen.data(n, p, family = "gaussian", K, rho, sigma)

# Best subset selection
fit1 <- bess.one(data$x, data$y, s = 10, family = "gaussian", normalize = TRUE)
#coef(fit1,sparse=TRUE)
bestmodel <- fit1$bestmodel
#summary(bestmodel)

## Not run:
#-----logistic model-----#

# Generate simulated data
data <- gen.data(n, p, family = "binomial", K, rho, sigma)

# Best subset selection
fit2 <- bess.one(data$x, data$y, family = "binomial", s = 10, normalize = TRUE)
bestmodel <- fit2$bestmodel
#summary(bestmodel)

#-----cox model-----#

# Generate simulated data
data <- gen.data(n, p, K, rho, sigma, c=10, family="cox", scal=10)

# Best subset selection
fit3 <- bess.one(data$x, data$y, s = 10, family = "cox", normalize = TRUE)
bestmodel <- fit3$bestmodel
#summary(bestmodel)

#-----High dimensional linear models-----#

p <- 1000
data <- gen.data(n, p, family = "gaussian", K, rho, sigma)

# Best subset selection
fit <- bess.one(data$x, data$y, s=10, family = "gaussian", normalize = TRUE)

#-----prostate-----#
data("prostate")
x = prostate[,-9]
y = prostate[,9]

fit.ungroup = bess.one(x, y, s=5)
fit.group = bess.one(x, y, s=5, factor = c("gleason"))

#-----SAheart-----#
data(SAheart)
y = SAheart[,5]

```

```
x = SAheart[,-5]
x$ldl[x$ldl<5] = 1
x$ldl[x$ldl>=5&x$ldl<10] = 2
x$ldl[x$ldl>=10] = 3

fit.ungroup = bess.one(x, y, s=5, family = "binomial")
fit.group = bess.one(x, y, s=5, factor = c("ldl"), family = "binomial")
## End(Not run)
```

coef.bess

Provides estimated coefficients from a fitted "bess" object.

Description

Similar to other prediction methods, this function provides estimated coefficients from a fitted "bess" object.

Usage

```
## S3 method for class 'bess'
coef(object, sparse=TRUE, type = c("ALL", "AIC", "BIC", "EBIC"),...)
```

Arguments

object	A "bess" project or a "bess.one" project.
sparse	Logical or NULL, specifying whether the coefficients should be presented as sparse matrix or not.
type	Types of coefficients returned. type = "AIC" cooresponds to the coefficient with optimal AIC value; type = "BIC" cooresponds to the coefficient with optimal BIC value; type = "EBIC" cooresponds to the coefficient with optimal EBIC value; type = "ALL" cooresponds to all coefficients in the bess object. Default is ALL.
...	Other arguments.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
coef(fit, sparse=TRUE) # The estimated coefficients
```

coef.bess.one	<i>Provides estimated coefficients from a fitted "bess.one" object.</i>
---------------	---

Description

Similar to other prediction methods, this function provides estimated coefficients from a fitted "bess.one" object.

Usage

```
## S3 method for class 'bess.one'
coef(object, sparse = TRUE, ...)
```

Arguments

object	A "bess.one" project.
sparse	Logical or NULL, specifying whether the coefficients should be presented as sparse matrix or not.
...	Other arguments.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")
coef(fit, sparse=TRUE) # The estimated coefficients
```

deviance.bess *Extract the deviance from a "bess" object.*

Description

Similar to other deviance methods, which returns deviance from a fitted "bess" object.

Usage

```
## S3 method for class 'bess'  
deviance(object, ...)
```

Arguments

object Output from the bess function or the bess.one function.
... Additional arguments affecting the predictions produced.

Value

The value of the deviance extracted from the "bess" object.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)  
fit <- bess(data$x, data$y, family = "gaussian")  
deviance(fit)
```

deviance.bess.one *Extract the deviance from a "bess.one" object.*

Description

Similar to other deviance methods, which returns deviance from a fitted "bess.one" object.

Usage

```
## S3 method for class 'bess.one'  
deviance(object,...)
```

Arguments

object Output from the `bess.one` function.
... Additional arguments affecting the predictions produced.

Value

The value of the deviance extracted from the "bess.one" object.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)  
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")  
deviance(fit)
```

gen.data *Generate simulated data*

Description

Generate data for simulations under the generalized linear model and Cox model.

Usage

```
gen.data(n, p, family, K, rho = 0, sigma = 1, beta = NULL, censoring = TRUE,
         c = 1, scal)
```

Arguments

n	The number of observations.
p	The number of predictors of interest.
family	The distribution of the simulated data. "gaussian" for gaussian data. "binomial" for binary data. "cox" for survival data
K	The number of nonzero coefficients in the underlying regression model.
rho	A parameter used to characterize the pairwise correlation in predictors. Default is 0.
sigma	A parameter used to control the signal-to-noise ratio. For linear regression, it is the error variance σ^2 . For logistic regression and Cox's model, the larger the value of sigma, the higher the signal-to-noise ratio.
beta	The coefficient values in the underlying regression model.
censoring	Whether data is censored or not. Default is TRUE
c	The censoring rate. Default is 1.
scal	A parameter in generating survival time based on the Weibull distribution. Only used for the "cox" family.

Details

For the design matrix X , we first generate an $n \times p$ random Gaussian matrix \bar{X} whose entries are i.i.d. $\sim N(0, 1)$ and then normalize its columns to the \sqrt{n} length. Then the design matrix X is generated with $X_j = \bar{X}_j + \rho(\bar{X}_{j+1} + \bar{X}_{j-1})$ for $j = 2, \dots, p-1$.

For "gaussian" family, the data model is

$$Y = X\beta + \epsilon, \text{ where } \epsilon \sim N(0, \sigma^2).$$

The underlying regression coefficient β has uniform distribution $[m, 100m]$, $m = 5\sqrt{2\log(p)/n}$.

For "binomial" family, the data model is

$$\text{Prob}(Y = 1) = \exp(X\beta)/(1 + \exp(X\beta))$$

The underlying regression coefficient β has uniform distribution $[2m, 10m]$, $m = 5\sigma\sqrt{2\log(p)/n}$.

For "cox" family, the data model is

$$T = (-\log(S(t))/\exp(X\beta))^{1/\text{scal}},$$

The centering time C is generated from uniform distribution $[0, c]$, then we define the censor status as $\delta = IT \leq C, R = \min T, C$. The underlying regression coefficient β has uniform distribution $[2m, 10m]$, $m = 5\sigma\sqrt{2\log(p)/n}$.

Value

A list with the following components: x, y, Tbeta.

x	Design matrix of predictors.
y	Response variable
Tbeta	The coefficients used in the underlying regression model.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

Examples

```
# Generate simulated data
n <- 500
p <- 20
K <- 10
sigma <- 1
rho <- 0.2
data <- gen.data(n, p, family = "gaussian", K, rho, sigma)

# Best subset selection
fit <- bess(data$x, data$y, family = "gaussian")
```

gravier	<i>breast cancer data set</i>
---------	-------------------------------

Description

Gravier et al. (2010) have considered small, invasive ductal carcinomas without axillary lymph node involvement (T1T2N0) to predict metastasis of small node-negative breast carcinoma. Using comparative genomic hybridization arrays, they examined 168 patients over a five-year period. The 111 patients with no event after diagnosis were labelled good, and the 57 patients with early metastasis were labelled poor.

Usage

```
data(gravier)
```

Format

A list containing the design matrix X and response matrix y

Source

<https://github.com/ramhiser>

References

Eleonore Gravier., Gaelle Pierron., and Anne Vincent-Salomon. (2010). A prognostic DNA signature for T1T2 node-negative breast cancer patients.

logLik.bess	<i>Extract the loglikelihood from a "bess" object.</i>
-------------	--

Description

Similar to other logLik methods, which returns loglikelihood from a fitted "bess" object.

Usage

```
## S3 method for class 'bess'  
logLik(object,...)
```

Arguments

object	Output from the bess function.
...	Additional arguments affecting the predictions produced.

Value

The value of the loglikelihood extracted from the "bess" object.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
logLik(fit)
```

logLik.bess.one	<i>Extract the loglikelihood from a "bess.one" object.</i>
-----------------	--

Description

Similar to other logLik methods, which returns loglikelihood from a fitted "bess.one" object.

Usage

```
## S3 method for class 'bess.one'
logLik(object,...)
```

Arguments

object	Output from the bess.one function.
...	Additional arguments affecting the predictions produced.

Value

The value of the loglikelihood extracted from the "bess.one" object.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")
logLik(fit)
```

plot.bess	<i>Produces a coefficient profile plot of the coefficient or loss function paths</i>
-----------	--

Description

Produces a coefficient profile plot of the coefficient or loss paths for a fitted "bess" object.

Usage

```
## S3 method for class 'bess'
plot(x, type=c("loss", "coefficients", "both"), breaks=TRUE, K=NULL, ...)
```

Arguments

x	a "bess" project
type	Either "both", "solutionPath" or "loss"
breaks	If TRUE, then vertical lines are drawn at each break point in the coefficient paths
K	which break point should the vertical lines drawn at
...	Other graphical parameters to plot

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also[bess](#), [bess.one](#)**Examples**

```
#-----linear model-----#
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
plot(fit, type = "both")
```

predict.bess	<i>make predictions from a "bess" object.</i>
--------------	---

Description

Similar to other predict methods, which returns predictions from a fitted "bess" object.

Usage

```
## S3 method for class 'bess'
predict(object, newdata, type = c("ALL", "opt", "AIC", "BIC", "EBIC"),...)
```

Arguments

object	Output from the bess function or the bess.one function.
newdata	New data used for prediction.
type	Types of coefficients returned. type = "AIC" cooresponds to the predictor with optimal AIC value; type = "BIC" cooresponds to the predictor with optimal BIC value; type = "EBIC" cooresponds to the predictor with optimal EBIC value; type = "ALL" cooresponds to all predictors in the bess object; type = "opt" cooresponds to predictors in best model. Default is ALL.
...	Additional arguments affecting the predictions produced.

Value

The object returned depends on the types of family.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
pred=predict(fit, newdata = data$x)
```

predict.bess.one	<i>make predictions from a "bess.one" object.</i>
------------------	---

Description

Similar to other predict methods, which returns predictions from a fitted "bess.one" object.

Usage

```
## S3 method for class 'bess.one'
predict(object, newdata, ...)
```

Arguments

object	Output from the <code>bess.one</code> function.
newdata	New data used for prediction.
...	Additional arguments affecting the predictions produced.

Value

The object returned depends on the types of family.

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")
pred <- predict(fit, newdata = data$x)
```

print.bess	<i>print method for a "bess" object</i>
------------	---

Description

Print the primary elements of the "bess" object.

Usage

```
## S3 method for class 'bess'
print(x, ...)
```

Arguments

x	a "bess" object
...	additional print arguments

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
print(fit)
```

print.bess.one	<i>print method for a "bess.one" object</i>
----------------	---

Description

Print the primary elements of the "bess.one" object.

Usage

```
## S3 method for class 'bess.one'  
print(x, ...)
```

Arguments

x	a "bess.one" object
...	additional print arguments

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)  
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")  
print(fit)
```

prostate

Factors associated with prostate specific antigen

Description

Data from a study by Stamey et al. (1989) to examine the association between prostate specific antigen (PSA) and several clinical measures that are potentially associated with PSA in men who were about to receive a radical prostatectomy. The variables are as follows:

- `lcavol`: Log cancer volume
- `lweight`: Log prostate weight
- `age`: The man's age
- `lbph`: Log of the amount of benign hyperplasia
- `svi`: Seminal vesicle invasion; 1=Yes, 0=No
- `lcp`: Log of capsular penetration
- `gleason`: Gleason score
- `pgg45`: Percent of Gleason scores 4 or 5
- `lpsa`: Log PSA

Usage

```
data(prostate)
```

Format

A data frame with 97 observations on 9 variables

References

Stamey, T., Kabalin, J., McNeal, J., Johnstone, I., Freiha, F., Redwine, E. and Yang, N. (1989). Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate II. Radical prostatectomy treated patients, *Journal of Urology* 16: 1076-1083.

SAheart

Risk factors associated with heart disease

Description

Data from a subset of the Coronary Risk-Factor Study baseline survey, carried out in rural South Africa. The variables are as follows:

- sbp: Systolic blood pressure
- tobacco: Cumulative tobacco consumption, in kg
- ldl: Low-density lipoprotein cholesterol
- adiposity: Adipose tissue concentration
- famhist: Family history of heart disease (1=Present, 0=Absent)
- typea: Score on test designed to measure type-A behavior
- obesity: Obesity
- alcohol: Current consumption of alcohol
- age: Age of subject
- chd: Coronary heart disease at baseline; 1=Yes 0=No

Usage

```
data(SAheart)
```

Format

A data frame with 462 observations on 10 variables

References

Rousseauw, J., du Plessis, J., Benade, A., Jordaan, P., Kotze, J. and Ferreira, J. (1983). Coronary risk factor screening in three rural communities. South African Medical Journal 64: 430-436.

summary.bess

summary method for a "bess" object

Description

Print a summary of the "bess" object.

Usage

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


Arguments

object a "bess" object
... additional print arguments

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess(data$x, data$y, family = "gaussian")
summary(fit)
```

summary.bess.one *summary method for a "bess.one" object*

Description

Print a summary of the "bess.one" object.

Usage

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

Arguments

object a "bess.one" object
... additional print arguments

Author(s)

Canhong Wen, Aijun Zhang, Shijie Quan, and Xueqin Wang.

References

Wen, C., Zhang, A., Quan, S. and Wang, X. (2020). BeSS: An R Package for Best Subset Selection in Linear, Logistic and Cox Proportional Hazards Models, *Journal of Statistical Software*, Vol. 94(4). doi:10.18637/jss.v094.i04.

See Also

[bess](#), [bess.one](#)

Examples

```
data <- gen.data(500, 20, family = "gaussian", 10, 0.2, 1)
fit <- bess.one(data$x, data$y, s = 10, family = "gaussian")
summary(fit)
```

Index

* datasets

prostate, [23](#)

SAheart, [24](#)

aic, [2](#)

bess, [3](#), [3](#), [8](#), [10–13](#), [17–22](#), [25](#), [26](#)

bess.one, [3](#), [5](#), [7](#), [10–13](#), [17–22](#), [25](#), [26](#)

bic(aic), [2](#)

coef.bess, [10](#)

coef.bess.one, [11](#)

deviance.bess, [12](#)

deviance.bess.one, [13](#)

ebic(aic), [2](#)

gen.data, [14](#)

gravier, [16](#)

logLik.bess, [16](#)

logLik.bess.one, [17](#)

plot.bess, [5](#), [8](#), [18](#)

predict.bess, [5](#), [8](#), [19](#)

predict.bess.one, [20](#)

print.bess, [21](#)

print.bess.one, [22](#)

prostate, [23](#)

SAheart, [24](#)

summary.bess, [24](#)

summary.bess.one, [25](#)