Package ‘ordinalgmifs’

July 16, 2018

Version 1.0.5
Date 2018-07-12
Title Ordinal Regression for High-Dimensional Data
Author Kellie J. Archer, Jiayi Hou, Qing Zhou, Kyle Ferber, John G. Layne, Amanda Gentry
Maintainer Kellie J. Archer <archer.43@osu.edu>
Depends R (>= 2.10), survival
Description Provides a function for fitting cumulative link, adjacent category, forward and backward continuation ratio, and stereotype ordinal response models when the number of parameters exceeds the sample size, using the the generalized monotone incremental forward stagewise method.
License GPL (>= 2)
BuildResaveData best
SystemRequirements C++11
NeedsCompilation yes
BuildVignettes TRUE
LazyData true
Repository CRAN
Date/Publication 2018-07-16 15:20:06 UTC

R topics documented:

  ordinalgmifs-package .................................................. 2
  coef.ordinalgmifs ..................................................... 3
  eyedisease .............................................................. 4
  hccframe ............................................................... 5
  ordinalgmifs ........................................................... 7
  plot.ordinalgmifs ...................................................... 9
  predict.ordinalgmifs .................................................. 10
  print.ordinalgmifs .................................................... 11
  summary.ordinalgmifs ................................................ 12

Index 13
Description

This package provides a function, ordinalgmifs, for fitting cumulative link, adjacent category, forward and backward continuation ratio, and stereotype ordinal response models when the number of parameters exceeds the sample size, using the the generalized monotone incremental forward stagewise method.

Details

Package: ordinalgmifs
Version: 1.0.5
Date: 2018-07-12
Title: Ordinal Regression for High-Dimensional Data
Author: Kellie J. Archer, Jiayi Hou, Qing Zhou, Kyle Ferber, John G. Layne, Amanda Gentry
Maintainer: Kellie J. Archer <archer.43@osu.edu>
Depends: R (>= 2.10), survival
Description: Provides a function for fitting cumulative link, adjacent category, forward and backward continuation ratio, and stereotype ordinal response models when the number of parameters exceeds the sample size, using the generalized monotone incremental forward stagewise method.
License: GPL (>=2)
BuildResaveData: best
SystemRequirements: C++11
NeedsCompilation: yes
BuildVignettes: TRUE
LazyData: true

Index of help topics:

- coef.ordinalgmifs: Extract Model Coefficients
- eyedisease: Eye Disease Risk Factors
- hccframe: Liver Cancer Methylation Data
- ordinalgmifs: Ordinal Generalized Monotone Incremental Forward Stagewise Regression
- ordinalgmifs-package: Ordinal Response Regression for High-Dimensional Data
- plot.ordinalgmifs: Plot Solution Path for Ordinal GMIFS Fitted Model.
- predict.ordinalgmifs: Predicted Probabilities and Class for Ordinal GMIFS Fit.
- print.ordinalgmifs: Print the Contents of an Ordinal GMIFS Fitted Object.
- summary.ordinalgmifs: Summarize an Ordinal GMIFS Object.

This package contains generic methods (coef, plot, predict, print, summary) that can be invoked for an object fitted using ordinalgmifs.
Author(s)
Kellie J. Archer, Jiayi Hou, Qing Zhou, Kyle Ferber, John G. Layne, Amanda Gentry
Maintainer: Kellie J. Archer <archer.43@osu.edu>

References

See Also
See Also ordinalgmifs. For models where no predictor is penalized see vglm

c coef. ordinalgmifs  Extract Model Coefficients

Description
coef. ordinalgmifs is a generic function which extracts the model coefficients from a fitted model
object fit using ordinalgmifs

Usage
## S3 method for class 'ordinalgmifs'
coef(object, model.select = "AIC", ...)

Arguments

object
an ordinalgmifs object.

model.select
when x is specified any model along the solution path can be selected. The
default is model.select="AIC" which extracts the coefficients from the model
having the lowest AIC. Other options are model.select="BIC" or any numeric
value from the solution path.

... other arguments.

Value
Coefficients extracted from the model object.

Author(s)
Kellie J. Archer

References

See Also

See Also `ordinalgmifs`, `summary.ordinalgmifs`, `plot.ordinalgmifs`, `predict.ordinalgmifs`

---

eyedisease  

**Eye Disease Risk Factors**

**Description**

Eye Disease Risk Factors data from Section 9.1 of Agresti’s Analysis of Ordinal Categorical Data. The primary data are from the Wisconsin Epidemiological Study of Diabetic Retinopathy. The primary outcome is severity of retinopathy which was measured in the left and right eye of every subject.

**Usage**

```r
data(eyedisease)
```

**Format**

A data frame with 720 observations on the following 19 variables.

- `rme` right eye macular oedema (absent = 0, present = 1)
- `lme` left eye macular oedema (absent = 0, present = 1)
- `rre` right eye refraction index
- `lre` left eye refraction index
- `riop` right eye intraocular eye pressure
- `liop` left eye intraocular eye pressure
- `age` age
- `diab` duration of diabetes (in years)
- `gh` glycosylated haemoglobin level
- `sbp` systolic blood pressure
- `dbp` diastolic blood pressure
- `bmi` body mass index
- `pr` pulse rate?
- `sex` gender (male=1, female=2)
- `prot` proteinuria (absent = 0, present = 1)
- `dose` a numeric vector
- `rerl` right eye severity of retinopathy, an ordered factor with levels `None` < `Mild` < `Moderate` < `Proliferative`
- `lerl` left eye severity of retinopathy, an ordered factor with levels `None` < `Mild` < `Moderate` < `Proliferative`
- `id` subject identifier
Description

These data are a subset of subjects and CpG sites reported in the original paper where liver samples were assayed using the Illumina GoldenGate Methylation BeadArray Cancer Panel I. Technical replicate samples were removed to ensure all samples were independent. The matched cirrhotic samples from subjects with hepatocellular carcinoma (HCC, labeled Tumor) were also excluded. Therefore methylation levels in liver tissue are provided for independent subjects whose liver was Normal (N=20), cirrhotic but not having HCC (N=16, Cirrhosis non-HCC), and HCC (N=20, Tumor).

Usage

data(hccframe)

Format

A data frame with 56 observations on the following 46 variables.

- **group**: an ordered factor with levels Normal < Cirrhosis non-HCC < Tumor
- **CDKN2B_seq_50_S294_F**: a numeric vector representing a CpG site proportion methylation for CDKN2B
- **DDIT3_P1313_R**: a numeric vector representing a CpG site proportion methylation for DDIT3
- **ERN1_P809_R**: a numeric vector representing a CpG site proportion methylation for ERN1
- **GML_E144_F**: a numeric vector representing a CpG site proportion methylation for GML

Source

See http://users.stat.ufl.edu/~aa/ordinal/data.html

References


See Also

See Also as `ordinalgmifs`

Examples

data(eyedisease)
HDAC9_P137_R  a numeric vector representing a CpG site proportion methylation for HDAC9
HLA_DPA1_P205_R  a numeric vector representing a CpG site proportion methylation for HLA.DPA1
HOXB2_P488_R  a numeric vector representing a CpG site proportion methylation for HOXB2
IL16_P226_F  a numeric vector representing a CpG site proportion methylation for IL16
IL16_P93_R  a numeric vector representing a CpG site proportion methylation for IL16
IL8_P83_F  a numeric vector representing a CpG site proportion methylation for IL8
MPO_E302_R  a numeric vector representing a CpG site proportion methylation for MPO
MPO_P883_R  a numeric vector representing a CpG site proportion methylation for MPO
PADI4_P1158_R  a numeric vector representing a CpG site proportion methylation for PADI4
SOX17_P287_R  a numeric vector representing a CpG site proportion methylation for SOX17
TJP2_P518_F  a numeric vector representing a CpG site proportion methylation for TJP2
WRN_E57_F  a numeric vector representing a CpG site proportion methylation for WRN
CRIP1_P874_R  a numeric vector representing a CpG site proportion methylation for CRIP1
SLC22A3_P634_F  a numeric vector representing a CpG site proportion methylation for SLC22A3
CCNA1_P216_F  a numeric vector representing a CpG site proportion methylation for CCNA1
SEPT9_P374_F  a numeric vector representing a CpG site proportion methylation for SEPT9
ITGA2_E120_F  a numeric vector representing a CpG site proportion methylation for ITGA2
ITGA6_P718_R  a numeric vector representing a CpG site proportion methylation for ITGA6
HGF_P1293_R  a numeric vector representing a CpG site proportion methylation for HGF
DLG3_E340_F  a numeric vector representing a CpG site proportion methylation for DLG3
APP_E8_F  a numeric vector representing a CpG site proportion methylation for APP
SFTPB_P689_R  a numeric vector representing a CpG site proportion methylation for SFTPB
PENK_P447_R  a numeric vector representing a CpG site proportion methylation for PENK
COMT_E401_F  a numeric vector representing a CpG site proportion methylation for COMT
NOTCH1_E452_R  a numeric vector representing a CpG site proportion methylation for NOTCH1
EPHA8_P456_R  a numeric vector representing a CpG site proportion methylation for EPHA8
WT1_P853_F  a numeric vector representing a CpG site proportion methylation for WT1
KLK10_P268_R  a numeric vector representing a CpG site proportion methylation for KLK10
PCDH1_P264_F  a numeric vector representing a CpG site proportion methylation for PCDH1
TDGF1_P428_R  a numeric vector representing a CpG site proportion methylation for TDGF1
EFNB3_P442_R  a numeric vector representing a CpG site proportion methylation for EFNB3
MMP19_P306_F  a numeric vector representing a CpG site proportion methylation for MMP19
FGFR2_P460_R  a numeric vector representing a CpG site proportion methylation for FGFR2
RAF1_P338_F  a numeric vector representing a CpG site proportion methylation for RAF1
BMPR2_E435_F  a numeric vector representing a CpG site proportion methylation for BMPR2
GRB10_P496_R  a numeric vector representing a CpG site proportion methylation for GRB10
CTSH_P238_F  a numeric vector representing a CpG site proportion methylation for CTSH
SLC6A8_seq_28_S227_F  a numeric vector representing a CpG site proportion methylation for SLC6A8
PLXDC1_P236_F  a numeric vector representing a CpG site proportion methylation for PLXDC1
TFE3_P421_F  a numeric vector representing a CpG site proportion methylation for TFE3
TSG101_P139_R  a numeric vector representing a CpG site proportion methylation for TSG101
Source

The full dataset is available as GSE18081 from Gene Expression Omnibus at http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE18081

References


See Also

See Also as ordinalgmifs

Examples

data(hccframe)

 ordinalgmifs

 Ordinal Generalized Monotone Incremental Forward Stagewise Regression

Description

This function can fit a cumulative link, adjacent category, forward and backward continuation ratio, and stereotype ordinal response model when the number of parameters exceeds the sample size, using the generalized monotone incremental forward stagewise method.

Usage

ordinalgmifs(formula, data, x = NULL, subset, epsilon = 0.001, tol = 1e-05, scale = TRUE, probability.model = "cumulative", link = "logit", verbose=FALSE, assumption=NULL, ...)

Arguments

formula an object of class "formula" (or one that can be coerced to that class): a symbolic description of the model to be fitted. The left side of the formula is the ordinal outcome while the variables on the right side of the formula are the covariates that are not included in the penalization process. Note that if all variables in the model are to be penalized, an intercept only model formula should be specified.

data an optional data frame, list or environment (or object coercible by as.data.frame to a data frame) containing the variables in the model.

x an optional matrix of predictors that are to be penalized in the model fitting process.
subset an optional vector specifying a subset of observations to be used in the fitting process.

epsilon small incremental amount used to update a coefficient at a given step.

tol the iterative process stops when the difference between successive log-likelihoods is less than this specified level of tolerance.

scale logical, if TRUE the penalized predictors are centered and scaled.

probability.model the type of ordinal response model to be fit. Can be "Cumulative", "AdjCategory", "ForwardCR", "BackwardCR", or "Stereotype"

link the link function used. Allowable links for "Cumulative", "ForwardCR", and "BackwardCR" are "logit", "probit", and "cloglog". For an "AdjCategory" model only a "logit" link is allowed; for a "Stereotype" model only a "logit" link is allowed.

verbose logical, if TRUE the step number is printed to the console (default is FALSE).

assumption integer, only use with probability.model = "ForwardCR" and link = "cloglog" to denote the assumption to use for discrete censored survival modeling. If assumption = 1, assume the observation was censored at the end of the discrete time interval in which the censoring occurred; if assumption = 2, assume the observation was censored at the beginning of the interval in which censoring occurred; if assumption = 3, assume constant hazard rate within the interval in which the censoring occurred; if no censoring occurs, do not specify a value for assumption.

... additional arguments

Details

A model specified as response~terms, x=penalized.terms where response is the ordinal response vector and terms is the series of variables in the model that are not to be penalized and x is a matrix of variables that are to be penalized. For example, terms may include the variables age and gender while x includes hundreds to thousands of features from a high-throughput genomic experiment. In the event that no baseline demographic/clinical characteristics/subject level variables are available or needed in terms (all variables are to be penalized) then the model is specified as response~1, x=penalized.terms.

Value

AIC a vector of AIC values for each step (if x is specified).

BIC a vector of BIC values for each step (if x is specified).

alpha the ordinal threshold estimates for the fitted model.

theta the coefficient estimates for the unpenalized variables (if terms are specified on the right hand side of the model formula).

beta the coefficient estimates for the penalized variables (if x is specified in the model).

phi the scaling coefficient estimates (if a "Stereotype" logit model is fit).
plot.ordinalgmifs

logLik  a vector of log-likelihood values for each step (if terms are specified on the right hand side of the model formula).
link  the link function used in the model fit.
model.select  the step at which the minimum AIC was observed (if terms are specified on the right hand side of the model formula).
probability.model  the model fit.
scale  logical indicating whether penalized variables were centered and scaled.
w  the unpenalized variables in the model (if any).
x  the penalized variables in the model (if any).
y  the ordinal response.

Author(s)

Kellie J. Archer, Jiayi Hou, Qing Zhou, Kyle Ferber, John G. Layne, Amanda Gentry

References


See Also

See Also coef.ordinalgmifs, summary.ordinalgmifs, plot.ordinalgmifs, predict.ordinalgmifs

Examples

data(hccframe)
# To minimize processing time, MPO_E302_R is coerced into the model and only a subset of
# two CpG sites (DDIT3_P1313_R and HDAC9_P137_R) are included as penalized covariates
# in this demonstration, and epsilon is set to 0.01
hcc.fit <- ordinalgmifs(group ~ MPO_E302_R, x = c("DDIT3_P1313_R", "HDAC9_P137_R"),
data = hccframe, epsilon = 0.01)
coef(hcc.fit)
summary(hcc.fit)
phat <- predict(hcc.fit)
head(phat$predicted)
table(phat$class, hccframe$group)

plot.ordinalgmifs  Plot Solution Path for Ordinal GMIFS Fitted Model.

Description

This function plots either the coefficient path, the AIC, or the log-likelihood for a fitted ordinalgmifs object.
predict.ordinalgmifs

Usage

```r
## S3 method for class 'ordinalgmifs'
plot(x, type = "trace", xlab=NULL, ylab=NULL, main=NULL, ...)
```

Arguments

- `x`: An `ordinalgmifs` object.
- `type`: Default is "trace" which plots the coefficient path for the fitted object. Also available are "AIC", "BIC", and "logLik".
- `xlab`: A default x-axis label will be used which can be changed by specifying a user-defined x-axis label.
- `ylab`: A default y-axis label will be used which can be changed by specifying a user-defined y-axis label.
- `main`: A default main title will be used which can be changed by specifying a user-defined main title.
- `...`: Other arguments.

Author(s)

Kellie J. Archer

See Also

See Also `ordinalgmifs`, `coef.ordinalgmifs`, `summary.ordinalgmifs`, `predict.ordinalgmifs`

---

predict.ordinalgmifs

Predicted Probabilities and Class for Ordinal GMIFS Fit.

Description

This function returns a list that includes the predicted probabilities as well as the predicted class for an `ordinalgmifs` fitted object.

Usage

```r
## S3 method for class 'ordinalgmifs'
predict(object, neww = NULL, newdata, newx = NULL, model.select = "AIC", ...)
```

Arguments

- `object`: An `ordinalgmifs` fitted object.
- `neww`: An optional formula that includes the unpenalized variables to use for predicting the response. If omitted, the training data are used.
- `newdata`: An optional data.frame that minimally includes the unpenalized variables to use for predicting the response. If omitted, the training data are used.
print.ordinalgmifs

newx an optional matrix of penalized variables to use for predicting the response. If omitted, the training data are used.
model.select when x is specified any model along the solution path can be selected. The default is model.select="AIC" which calculates the predicted values using the coefficients from the model having the lowest AIC. Other options are model.select="BIC" or any numeric value from the solution path.

Value

predicted a matrix of predicted probabilities from the fitted model.
class a vector containing the predicted class taken as that class having the largest predicted probability.

Author(s)

Kellie J. Archer, Jiayi Hou, Qing Zhou, Kyle Ferber, John G. Layne, Amanda Gentry

See Also

See Also ordinalgmifs, coef.ordinalgmifs, summary.ordinalgmifs, plot.ordinalgmifs

Description

This function prints the names of the list objects from an ordinalgmifs fitted model.

Usage

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

Arguments

x an ordinalgmifs object.

Note

The contents of an ordinalgmifs fitted object differ depending upon whether x is specified in the ordinalgmifs model (i.e., penalized variables are included in the model fit hence a solution path is returned) or only terms on the right hand side of the equation are included (unpenalized variables). In the latter case, we recommend using the VGAM package.
Author(s)

Kellie J. Archer

See Also

See Also `ordinalgmifs`, `coef.ordinalgmifs`, `summary.ordinalgmifs`, `plot.ordinalgmifs`, `predict.ordinalgmifs`

summary.ordinalgmifs

Summarize an Ordinal GMIFS Object.

Description

summary method for class `ordinalgmifs`.

Usage

```r
## S3 method for class 'ordinalgmifs'
summary(object, model.select = "AIC", ...)
```

Arguments

- `object`: an `ordinalgmifs` object.
- `model.select`: when `x` is specified any model along the solution path can be selected. The default is `model.select="AIC"` which extracts the model having the lowest AIC. Other options are `model.select="BIC"` or any numeric value from the solution path.
- `...`: other arguments.

Details

Prints the following items extracted from the fitted `ordinalgmifs` object: the probability model and link used and model parameter estimates. For models that include `x`, the parameter estimates, AIC, BIC, and log-likelihood are printed for indicated `model.select` step or if `model.select` is not supplied the step at which the minimum AIC was observed.

Author(s)

Kellie J. Archer

See Also

See Also `ordinalgmifs`, `coef.ordinalgmifs`, `plot.ordinalgmifs`, `predict.ordinalgmifs`
Index

*Topic **classif**
  ordalgmifs-package, 2

*Topic **datasets**
  eyedisease, 4
  hccframe, 5

*Topic **methods**
  coef.ordinalgmifs, 3
  plot.ordinalgmifs, 9
  predict.ordinalgmifs, 10
  print.ordinalgmifs, 11
  summary.ordinalgmifs, 12

*Topic **models**
  ordinalgmifs, 7

*Topic **package**
  ordinalgmifs-package, 2

*Topic **regression**
  ordinalgmifs, 7

  coef.ordinalgmifs, 3, 9–12

  eyedisease, 4

  fitted.ordinalgmifs
    (predict.ordinalgmifs), 10

  hccframe, 5

  ordinalgmifs, 3–5, 7, 7, 10–12
  ordinalgmifs-package, 2

  plot.ordinalgmifs, 4, 9, 9, 11, 12
  predict.ordinalgmifs, 4, 9, 10, 10, 12
  print.ordinalgmifs, 11

  summary.ordinalgmifs, 4, 9–12, 12

  vglm, 3