Package ‘SLOPE’

December 10, 2021

Title Sorted L1 Penalized Estimation

Version 0.4.0

Description Efficient implementations for Sorted L-One Penalized Estimation (SLOPE): generalized linear models regularized with the sorted L1-norm (Bogdan et al. (2015) <doi:10/gfgwzt>). Supported models include ordinary least-squares regression, binomial regression, multinomial regression, and Poisson regression. Both dense and sparse predictor matrices are supported. In addition, the package features predictor screening rules that enable fast and efficient solutions to high-dimensional problems.

License GPL-3

LazyData true

Depends R (>= 3.3.0)

Imports foreach, lattice, Matrix, methods, Rcpp

LinkingTo Rcpp, RcppArmadillo (>= 0.9.850.1.0)

Suggests caret, covr, glmnet, ggplot2, stringr, scales, tidyr, dplyr, bench, knitr, rmarkdown, spelling, testthat (>= 2.1.0)

RoxygenNote 7.1.2

Language en-US

Encoding UTF-8


BugReports https://github.com/jolars/SLOPE/issues

VignetteBuilder knitr

NeedsCompilation yes

Author Johan Larsson [aut, cre] (<https://orcid.org/0000-0002-4029-5945>), Jonas Wallin [aut] (<https://orcid.org/0000-0003-0381-6593>), Malgorzata Bogdan [aut], Ewout van den Berg [aut], Chiara Sabatti [aut], Emmanuel Candès [aut], Evan Patterson [aut],
Abalone

Description
This data set contains observations of abalones, the common name for any of a group of sea snails. The goal is to predict the age of an individual abalone given physical measurements such as sex, weight, and height.
Usage

abalone

Format

A list with two items representing 211 observations from 9 variables

sex  sex of abalone, 1 for female
infant  indicates that the person is an infant
length  longest shell measurement in mm
diameter  perpendicular to length in mm
height  height in mm including meat in shell
weight_whole  weight of entire abalone
weight_shucked  weight of meat
weight_viscera  weight of viscera
weight_shell  weight of shell
rings  rings. +1.5 gives the age in years

Details

Only a stratified sample of 211 rows of the original data set are used here.

Source


See Also

Other datasets: bodyfat, heart, student, wine

---

Description

The response (y) corresponds to estimates of percentage of body fat from application of Siri’s 1956 equation to measurements of underwater weighing, as well as age, weight, height, and a variety of body circumference measurements.

Usage

bodyfat
Format

A list with two items representing 252 observations from 14 variables

- **age**  age (years)
- **weight**  weight (lbs)
- **height**  height (inches)
- **neck**  neck circumference (cm)
- **chest**  chest circumference (cm)
- **abdomen**  abdomen circumference (cm)
- **hip**  hip circumference (cm)
- **thigh**  thigh circumference (cm)
- **knee**  knee circumference (cm)
- **ankle**  ankle circumference (cm)
- **biceps**  biceps circumference (cm)
- **forearm**  forearm circumference (cm)
- **wrist**  wrist circumference (cm)

Source

http://lib.stat.cmu.edu/datasets/bodyfat
https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/regression.html

See Also

Other datasets: abalone, heart, student, wine

caretSLOPE  
Model objects for model tuning with caret

Description

This function can be used in a call to `caret::train()` to enable model tuning using caret. Note that this function does not properly work with sparse feature matrices and standardization due to the way resampling is implemented in caret. So for these cases, please check out `trainSLOPE()` instead.

Usage

caretSLOPE()

Value

A model description list to be used in the `method` argument in `caret::train()`. 
**coef.SLOPE**

**Obtain coefficients**

**Description**

This function returns coefficients from a model fit by `SLOPE()`.

**Usage**

```r
## S3 method for class 'SLOPE'
coef(object, alpha = NULL, exact = FALSE, simplify = TRUE, sigma, ...)
```

**Arguments**

- `object`: an object of class 'SLOPE'.
- `alpha`: penalty parameter for SLOPE models; if NULL, the values used in the original fit will be used.
- `exact`: if TRUE and the given parameter values differ from those in the original fit, the model will be refit by calling `stats::update()` on the object with the new parameters. If FALSE, the predicted values will be based on interpolated coefficients from the original penalty path.
- `simplify`: if TRUE, `base::drop()` will be called before returning the coefficients to drop extraneous dimensions.
- `sigma`: deprecated. Please use `alpha` instead.
- `...`: arguments that are passed on to `stats::update()` (and therefore also to `SLOPE()`) if `exact = TRUE` and the given penalty is not in `object`.

**Details**

If `exact = FALSE` and `alpha` is not in `object`, then the returned coefficients will be approximated by linear interpolation. If coefficients from another type of penalty sequence (with a different `lambda`) are required, however, please use `SLOPE()` to refit the model.

**Value**

Coefficients from the model.

**See Also**

- `predict.SLOPE()`, `SLOPE()`
- Other SLOPE-methods: `deviance.SLOPE()`, `plot.SLOPE()`, `predict.SLOPE()`, `print.SLOPE()`, `score()`
Examples

```r
fit <- SLOPE(mtcars$mpg, mtcars$vs, path_length = 1)
coef(fit)
```

---

**deviance.SLOPE**  
*Model deviance*

### Description

Model deviance

### Usage

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

### Arguments

- `object`: an object of class 'SLOPE'.
- `...`: ignored

### Value

For Gaussian models this is twice the residual sums of squares. For all other models, two times the negative loglikelihood is returned.

### See Also

- `SLOPE()`
- Other SLOPE-methods: `coef.SLOPE()`, `plot.SLOPE()`, `predict.SLOPE()`, `print.SLOPE()`, `score()`

### Examples

```r
fit <- SLOPE(abalone$x, abalone$y, family = "poisson")
device(fit)
```
**Heart Disease**

Description

Diagnostic attributes of patients classified as having heart disease or not.

Usage

heart

Format

270 observations from 17 variables represented as a list consisting of a binary factor response vector \( y \), with levels 'absence' and 'presence' indicating the absence or presence of heart disease and \( x \): a sparse feature matrix of class 'dgCMatrix' with the following variables:

- **age** age
- **bp** diastolic blood pressure
- **chol** serum cholesterol in mg/dl
- **hr** maximum heart rate achieved
- **old_peak** ST depression induced by exercise relative to rest
- **vessels** the number of major blood vessels (0 to 3) that were colored by fluoroscopy
- **sex** sex of the participant: 0 for male, 1 for female
- **angina** a dummy variable indicating whether the person suffered angina-pectoris during exercise
- **glucose_high** indicates a fasting blood sugar over 120 mg/dl
- **cp_typical** typical angina
- **cp_atypical** atypical angina
- **cp_nonanginal** non-anginal pain
- **ecg_abnormal** indicates a ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)
- **ecg_estes** probable or definite left ventricular hypertrophy by Estes’ criteria
- **slope_flat** a flat ST curve during peak exercise
- **slope_downsloping** a downwards-sloping ST curve during peak exercise
- **thal_reversible** reversible defect
- **thal_fixed** fixed defect

Preprocessing

The original dataset contained 13 variables. The nominal of these were dummycoded, removing the first category. No precise information regarding variables chest_pain, thal and ecg could be found, which explains their obscure definitions here.
Source
https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html#heart

See Also
Other datasets: abalone, bodyfat, student, wine

plot.SLOPE

Plot coefficients

Description
Plot the fitted model’s regression coefficients along the regularization path.

Usage
## S3 method for class 'SLOPE'
plot(
  x,
  intercept = FALSE,
  x_variable = c("alpha", "deviance_ratio", "step"),
  ...,
)

Arguments

x an object of class "SLOPE"
intercept whether to plot the intercept
x_variable what to plot on the x axis. "alpha" plots the scaling parameter for the sequence, "deviance_ratio" plots the fraction of deviance explained, and "step" plots step number.

Value
An object of class "trellis", which will be plotted on the current device unless stored in a variable.

See Also
lattice::xyplot(), SLOPE(), plotDiagnostics()
Other SLOPE-methods: coef.SLOPE(), deviance.SLOPE(), predict.SLOPE(), print.SLOPE(), score()
Examples

```r
fit <- SLOPE(heart$x, heart$y)
plot(fit)
```

Description

Plot results from cross-validation

Usage

```r
## S3 method for class 'TrainedSLOPE'
plot(
  x,
  measure = c("auto", "mse", "mae", "deviance", "auc", "misclass"),
  plot_min = TRUE,
  ci_alpha = 0.2,
  ci_border = FALSE,
  ci_col = lattice::trellis.par.get("superpose.line")$col,
  ...
)
```

Arguments

- `x`: an object of class 'TrainedSLOPE', typically from a call to `trainSLOPE()`
- `measure`: any of the measures used in the call to `trainSLOPE()`. If `measure = "auto"` then deviance will be used for binomial and multinomial models, whilst mean-squared error will be used for Gaussian and Poisson models.
- `plot_min`: whether to mark the location of the penalty corresponding to the best prediction score
- `ci_alpha`: alpha (opacity) for fill in confidence limits
- `ci_border`: color (or flag to turn off and on) the border of the confidence limits
- `ci_col`: color for border of confidence limits
- `...`: other arguments that are passed on to `lattice::xyplot()`

Value

An object of class "trellis" is returned and, if used interactively, will most likely have its print function `lattice::print.trellis()` invoked, which draws the plot on the current display device.

See Also

- `trainSLOPE()`, `lattice::xyplot()`, `lattice::panel.xyplot()`
- Other model-tuning: `caretSLOPE()`, `trainSLOPE()`
Examples

# Cross-validation for a SLOPE binomial model
set.seed(123)
tune <- trainSLOPE(subset(mtcars, select = c("mpg", "drat", "wt")),
  mtcars$hp,
  q = c(0.1, 0.2),
  number = 10)
plot(tune, ci_col = "salmon", col = "black")

plotDiagnostics

Plot results from diagnostics collected during model fitting

Description

This function plots various diagnostics collected during the model fitting resulting from a call to SLOPE() provided that diagnostics = TRUE.

Usage

plotDiagnostics(
  object,
  ind = max(object$diagnostics$penalty),
  xvar = c("time", "iteration"),
  ...
)

Arguments

object an object of class "SLOPE".
ind either "last"
xvar what to place on the x axis. iteration plots each iteration, time plots the wall-clock time.
... other arguments that will be used to modify the call to lattice::xyplot()

Value

An object of class "trellis", which, unless stored in a variable, will be plotted when its default print() method is called.

See Also

SLOPE()

Examples

x <- SLOPE(abalone$x, abalone$y, diagnostics = TRUE)
plotDiagnostics(x)
predict.SLOPE

Generate predictions from SLOPE models

Description

Return predictions from models fit by SLOPE().

Usage

## S3 method for class 'SLOPE'
predict(object, x, alpha = NULL, type = "link", simplify = TRUE, sigma, ...)

## S3 method for class 'GaussianSLOPE'
predict(
  object,
  x,
  sigma = NULL,
  type = c("link", "response"),
  simplify = TRUE,
...
)

## S3 method for class 'BinomialSLOPE'
predict(
  object,
  x,
  sigma = NULL,
  type = c("link", "response", "class"),
  simplify = TRUE,
...
)

## S3 method for class 'PoissonSLOPE'
predict(
  object,
  x,
  sigma = NULL,
  type = c("link", "response"),
  exact = FALSE,
  simplify = TRUE,
...
)

## S3 method for class 'MultinomialSLOPE'
predict(
  object,
  x,
sigma = NULL,
type = c("link", "response", "class"),
exact = FALSE,
simplify = TRUE,
...)

Arguments

object an object of class "SLOPE", typically the result of a call to SLOPE()
x new data
alpha penalty parameter for SLOPE models; if NULL, the values used in the original fit will be used
type type of prediction; "link" returns the linear predictors, "response" returns the result of applying the link function, and "class" returns class predictions.
simplify if TRUE, base::drop() will be called before returning the coefficients to drop extraneous dimensions
sigma deprecated. Please use alpha instead.
... ignored and only here for method consistency
exact if TRUE and the given parameter values differ from those in the original fit, the model will be refit by calling stats::update() on the object with the new parameters. If FALSE, the predicted values will be based on interpolated coefficients from the original penalty path.

Value

Predictions from the model with scale determined by type.

See Also

stats::predict(), stats::predict.glm(), coef.SLOPE()

Other SLOPE-methods: coef.SLOPE(), deviance.SLOPE(), plot.SLOPE(), print.SLOPE(), score()

Examples

fit <- with(mtcars, SLOPE(cbind(mpg, hp), vs, family = "binomial"))
predict(fit, with(mtcars, cbind(mpg, hp)), type = "class")
**print.SLOPE**

Print results from SLOPE fit

---

### Description

Print results from SLOPE fit

### Usage

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

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

### Arguments

- `x` an object of class 'SLOPE' or 'TrainedSLOPE'
- `...` other arguments passed to `print()`

### Value

Prints output on the screen

### See Also

`SLOPE()`, `print.SLOPE()`

Other SLOPE-methods: `coef.SLOPE()`, `deviance.SLOPE()`, `plot.SLOPE()`, `predict.SLOPE()`, `score()`

### Examples

```r
fit <- SLOPE(wine$x, wine$y, family = "multinomial")
print(fit, digits = 1)
```

---

**regularizationWeights**

Generate Regularization (Penalty) Weights for SLOPE

---

### Description

This function generates sequences of regularizations weights for use in `SLOPE()` (or elsewhere).
Usage

regularizationWeights(
  n_lambda = 100,
  type = c("bh", "gaussian", "oscar", "lasso"),
  q = 0.2,
  theta1 = 1,
  theta2 = 0.5,
  n = NULL
)

Arguments

  n_lambda  The number of lambdas to generate. This should typically be equal to the number of predictors in your data set.
  type      The type of lambda sequence to use. See documentation for in SLOPE(), including that related to the lambda parameter in that function.
  q          parameter controlling the shape of the lambda sequence, with usage varying depending on the type of path used and has no effect is a custom lambda sequence is used. Must be greater than 1e-6 and smaller than 1.
  theta1     parameter controlling the shape of the lambda sequence when lambda == "OSCAR". This parameter basically sets the intercept for the lambda sequence and is equivalent to $\lambda_1$ in the original OSCAR formulation.
  theta2     parameter controlling the shape of the lambda sequence when lambda == "OSCAR". This parameter basically sets the slope for the lambda sequence and is equivalent to $\lambda_2$ in the original OSCAR formulation.
  n          The number of rows (observations) in the design matrix.

Details

Please see SLOPE() for detailed information regarding the parameters in this function, in particular the section Regularization Sequences.

Note that these sequences are automatically scaled (unless a value for the alpha parameter is manually supplied) when using SLOPE(). In this function, no such scaling is attempted.

Value

A vector of length n_lambda with regularization weights.

See Also

SLOPE()

Examples

# compute different penalization sequences
bh <- regularizationWeights(100, q = 0.2, type = "bh")
```
Score <- regularizationWeights(
    100,
    q = 0.2,
    n = 300,
    type = "gaussian"
)

oscar <- regularizationWeights(
    100,
    theta1 = 1.284,
    theta2 = 0.0182,
    type = "oscar"
)

lasso <- regularizationWeights(100, type = "lasso") * mean(oscar)

# Plot a comparison between these sequences
plot(bh, type = "l", ylab = expression(lambda))
lines(gaussian, col = "dark orange")
lines(oscar, col = "navy")
lines(lasso, col = "red3")
legend(
    "topright",
    legend = c("BH", "Gaussian", "OSCAR", "lasso"),
    col = c("black", "dark orange", "navy", "red3"),
    lty = 1
)
```

**score**

Compute one of several loss metrics on a new data set

**Description**

This function is a unified interface to return various types of loss for a model fit with SLOPE().

**Usage**

```r
score(object, x, y, measure)
```

```r
## S3 method for class 'GaussianSLOPE'
score(object, x, y, measure = c("mse", "mae"))
```

```r
## S3 method for class 'BinomialSLOPE'
score(object, x, y, measure = c("mse", "mae", "deviance", "misclass", "auc"))
```

```r
## S3 method for class 'MultinomialSLOPE'
score(object, x, y, measure = c("mse", "mae", "deviance", "misclass"))
```
## S3 method for class 'PoissonSLOPE'
score(object, x, y, measure = c("mse", "mae"))

### Arguments

- **object**: an object of class "SLOPE"
- **x**: feature matrix
- **y**: response
- **measure**: type of target measure. "mse" returns mean squared error. "mae" returns mean absolute error, "misclass" returns misclassification rate, and "auc" returns area under the ROC curve.

### Value

The measure along the regularization path depending on the value in `measure`.

### See Also

- `SLOPE()`, `predict.SLOPE()`

Other SLOPE-methods: `coef.SLOPE()`, `deviance.SLOPE()`, `plot.SLOPE()`, `predict.SLOPE()`, `print.SLOPE()`

### Examples

```r
x <- subset(infert, select = c("induced", "age", "pooled.stratum"))
y <- infert$case

fit <- SLOPE(x, y, family = "binomial")
score(fit, x, y, measure = "auc")
```

---

**SLOPE**  
Sort L-One Penalized Estimation

### Description

Fit a generalized linear model regularized with the sorted L1 norm, which applies a non-increasing regularization sequence to the coefficient vector ($\beta$) after having sorted it in decreasing order according to its absolute values.

### Usage

```r
SLOPE(
x,
y,
family = c("gaussian", "binomial", "multinomial", "poisson"),
intercept = TRUE,
center = !inherits(x, "sparseMatrix"),
```
scale = c("l2", "l1", "sd", "none"),
alph = c("path", "estimate"),
lambd = c("bh", "gaussian", "oscar", "lasso"),
alpha_min_ratio = if (NROW(x) < NCOL(x)) 0.01 else 1e-04,
path_length = if (alpha[1] == "estimate") 1 else 20,
q = 0.1 * min(1, NROW(x)/NCOL(x)),
theta1 = 1,
theta2 = 0.5,
prox_method = c("stack", "pava"),
screen = TRUE,
screen_alg = c("strong", "previous"),
tol_dev_change = 1e-05,
tol_dev_ratio = 0.995,
max_variables = NROW(x),
solver = c("fista", "admm"),
max_passes = 1e+06,
tol_abs = 1e-05,
tol_rel = 1e-04,
tol_rel_gap = 1e-05,
tol_infeas = 0.001,
tol_rel_coef_change = 0.001,
diagnostics = FALSE,
verbosity = 0,
theta1 = 1,
theta2 = 0.5,
prox_method = c("stack", "pava"),
screen = TRUE,
screen_alg = c("strong", "previous"),
tol_dev_change = 1e-05,
tol_dev_ratio = 0.995,
max_variables = NROW(x),
solver = c("fista", "admm"),
max_passes = 1e+06,
tol_abs = 1e-05,
tol_rel = 1e-04,
tol_rel_gap = 1e-05,
tol_infeas = 0.001,
tol_rel_coef_change = 0.001,
diagnostics = FALSE,
verbosity = 0,
sigma,
n_sigma,
lambda_min_ratio
)

Arguments

x
the design matrix, which can be either a dense matrix of the standard matrix class, or a sparse matrix inheriting from Matrix::sparseMatrix. Data frames will be converted to matrices internally.

y
the response, which for family = "gaussian" must be numeric; for family = "binomial" or family = "multinomial", it can be a factor.

family
model family (objective); see Families for details.

intercept
whether to fit an intercept

center
whether to center predictors or not by their mean. Defaults to TRUE if x is dense and FALSE otherwise.

scale
type of scaling to apply to predictors.

- "l1" scales predictors to have L1 norms of one.
- "l2" scales predictors to have L2 norms of one.
- "sd" scales predictors to have a population standard deviation one.
- "none" applies no scaling.

alpha
scale for regularization path: either a decreasing numeric vector (possibly of length 1) or a character vector; in the latter case, the choices are:
• "path", which computes a regularization sequence where the first value corresponds to the intercept-only (null) model and the last to the almost-saturated model, and
• "estimate", which estimates a single alpha using Algorithm 5 in Bogdan et al. (2015).

When a value is manually entered for alpha, it will be scaled based on the type of standardization that is applied to x. For scale = "l2", alpha will be scaled by \( \sqrt{n} \). For scale = "sd" or "none", alpha will be scaled by \( n \), and for scale = "l1" no scaling is applied. Note, however, that the alpha that is returned in the resulting value is the unstandardized alpha.

\textbf{lambda}

either a character vector indicating the method used to construct the lambda path or a numeric non-decreasing vector with length equal to the number of coefficients in the model; see section Regularization sequences for details.

\textbf{alpha_min_ratio}

smallest value for lambda as a fraction of lambda_max; used in the selection of alpha when alpha = "path".

\textbf{path_length}

length of regularization path; note that the path returned may still be shorter due to the early termination criteria given by tol_dev_change, tol_dev_ratio, and max_variables.

\textbf{q}

parameter controlling the shape of the lambda sequence, with usage varying depending on the type of path used and has no effect when a custom lambda sequence is used. Must be greater than 1e-6 and smaller than 1.

\textbf{theta1}

parameter controlling the shape of the lambda sequence when lambda == "OSCAR". This parameter basically sets the intercept for the lambda sequence and is equivalent to \( \lambda_1 \) in the original OSCAR formulation.

\textbf{theta2}

parameter controlling the shape of the lambda sequence when lambda == "OSCAR". This parameter basically sets the slope for the lambda sequence and is equivalent to \( \lambda_2 \) in the original OSCAR formulation.

\textbf{prox_method}

method for calculating the proximal operator for the Sorted L1 Norm (the SLOPE penalty). Please see sortedL1Prox() for more information.

\textbf{screen}

whether to use predictor screening rules (rules that allow some predictors to be discarded prior to fitting), which improve speed greatly when the number of predictors is larger than the number of observations.

\textbf{screen_alg}

what type of screening algorithm to use.

• "strong" uses the set from the strong screening rule and check against the full set
• "previous" first fits with the previous active set, then checks against the strong set, and finally against the full set if there are no violations in the strong set

\textbf{tol_dev_change}

the regularization path is stopped if the fractional change in deviance falls below this value; note that this is automatically set to 0 if a alpha is manually entered

\textbf{tol_dev_ratio}

the regularization path is stopped if the deviance ratio \( 1 - \text{deviance}/(\text{null} - \text{deviance}) \) is above this threshold
max_variables criterion for stopping the path in terms of the maximum number of unique, nonzero coefficients in absolute value in model. For the multinomial family, this value will be multiplied internally with the number of levels of the response minus one.
solver type of solver use, either "fista" or "admm"; all families currently support FISTA but only family = "gaussian" supports ADMM.
max_passes maximum number of passes (outer iterations) for solver
tol_abs absolute tolerance criterion for ADMM solver
tol_rel relative tolerance criterion for ADMM solver
tol_rel_gap stopping criterion for the duality gap; used only with FISTA solver.
tol_infeas stopping criterion for the level of infeasibility; used with FISTA solver and KKT checks in screening algorithm.
tol_rel_coef_change relative tolerance criterion for change in coefficients between iterations, which is reached when the maximum absolute change in any coefficient divided by the maximum absolute coefficient size is less than this value.
diagnostics whether to save diagnostics from the solver (timings and other values depending on type of solver)
verbosity level of verbosity for displaying output from the program. Setting this to 1 displays basic information on the path level, 2 a little bit more information on the path level, and 3 displays information from the solver.
sigma deprecated; please use alpha instead
n_sigma deprecated; please use path_length instead
lambda_min_ratio deprecated; please use alpha_min_ratio instead

Details

SLOPE() solves the convex minimization problem

\[ f(\beta) + \alpha \sum_{i=1}^{p} \lambda_i |\beta|_{(i)} \]

where \( f(\beta) \) is a smooth and convex function and the second part is the sorted L1-norm. In ordinary least-squares regression, \( f(\beta) \) is simply the squared norm of the least-squares residuals. See section Families for specifics regarding the various types of \( f(\beta) \) (model families) that are allowed in SLOPE().

By default, SLOPE() fits a path of models, each corresponding to a separate regularization sequence, starting from the null (intercept-only) model to an almost completely unregularized model. These regularization sequences are parameterized using \( \lambda \) and \( \alpha \), with only \( \alpha \) varying along the path. The length of the path can be manually, but will terminate prematurely depending on arguments tol_dev_change, tol_dev_ratio, and max_variables. This means that unless these arguments are modified, the path is not guaranteed to be of length path_length.
Value

An object of class "SLOPE" with the following slots:

- **coefficients**: a three-dimensional array of the coefficients from the model fit, including the intercept if it was fit. There is one row for each coefficient, one column for each target (dependent variable), and one slice for each penalty.
- **nonzeros**: a three-dimensional logical array indicating whether a coefficient was zero or not.
- **lambda**: the lambda vector that when multiplied by a value in alpha gives the penalty vector at that point along the regularization path.
- **alpha**: vector giving the (unstandardized) scaling of the lambda sequence.
- **class_names**: a character vector giving the names of the classes for binomial and multinomial families.
- **passes**: the number of passes the solver took at each step on the path.
- **violations**: the number of violations of the screening rule at each step on the path; only available if diagnostics = TRUE in the call to SLOPE().
- **active_sets**: a list where each element indicates the indices of the coefficients that were active at that point in the regularization path.
- **unique**: the number of unique predictors (in absolute value).
- **deviance_ratio**: the deviance ratio (as a fraction of 1).
- **null_deviance**: the deviance of the null (intercept-only) model.
- **family**: the name of the family used in the model fit.
- **diagnostics**: a data.frame of objective values for the primal and dual problems, as well as a measure of the infeasibility, time, and iteration; only available if diagnostics = TRUE in the call to SLOPE().
- **call**: the call used for fitting the model.

Families

**Gaussian**

The Gaussian model (Ordinary Least Squares) minimizes the following objective:

\[
\frac{1}{2} \| y - X \beta \|_2^2
\]

**Binomial**

The binomial model (logistic regression) has the following objective:

\[
\sum_{i=1}^{n} \log \left( 1 + \exp \left( -y_i \left( x_i^T \beta + \beta_0 \right) \right) \right)
\]

with \( y \in \{-1, 1\} \).

**Poisson**

In poisson regression, we use the following objective:
Multinomial Regression

In multinomial regression, we minimize the full-rank objective

\[- \sum_{i=1}^{n} \left( \sum_{k=1}^{m-1} y_{ik} (x_i^T \beta_k + \beta_{0,k}) - \log \sum_{k=1}^{m-1} \exp (x_i^T \beta_k + \beta_{0,k}) \right)\]

with $y_{ik}$ being the element in a $n \times (m-1)$ matrix, where $m$ is the number of classes in the response.

Regularization Sequences

There are multiple ways of specifying the lambda sequence in SLOPE(). It is, first of all, possible to select the sequence manually by using a non-increasing numeric vector, possibly of length one, as argument instead of a character. The greater the differences are between consecutive values along the sequence, the more clustering behavior will the model exhibit. Note, also, that the scale of the $\lambda$ vector makes no difference if $\alpha = \text{NULL}$, since $\alpha$ will be selected automatically to ensure that the model is completely sparse at the beginning and almost unregularized at the end. If, however, both $\alpha$ and $\lambda$ are manually specified, then the scales of both do matter, so make sure to choose them wisely.

Instead of choosing the sequence manually, one of the following automatically generated sequences may be chosen.

**BH (Benjamini–Hochberg)**

If $\lambda = \text{"bh"}$, the sequence used is that referred to as $\lambda^{(BH)}$ by Bogdan et al, which sets $\lambda$ according to

\[\lambda_i = \Phi^{-1} \left( 1 - iq / (2p) \right),\]

for $i = 1, \ldots, p$, where $\Phi^{-1}$ is the quantile function for the standard normal distribution and $q$ is a parameter that can be set by the user in the call to SLOPE().

**Gaussian**

This penalty sequence is related to BH, such that

\[\lambda_i = \lambda_i^{(BH)} \sqrt{1 + w(i-1) \cdot \text{cumsum}(\lambda^2)},\]

for $i = 1, \ldots, p$, where $w(k) = 1 / (n - k - 1)$. We let $\lambda_1 = \lambda_1^{(BH)}$ and adjust the sequence to make sure that it’s non-increasing. Note that if $p$ is large relative to $n$, this option will result in a constant sequence, which is usually not what you would want.

**OSCAR**

This sequence comes from Bondell and Reich and is a linear non-increasing sequence, such that

\[\lambda_i = \theta_1 + (p - i) \theta_2,\]

for $i = 1, \ldots, p$. We use the parametrization from Zhong and Kwok (2021) but use $\theta_1$ and $\theta_2$ instead of $\lambda_1$ and $\lambda_2$ to avoid confusion and abuse of notation.
SLOPE

SLOPE is exactly equivalent to the lasso when the sequence of regularization weights is constant, i.e.

$$\lambda_i = 1$$

for $i = 1, \ldots, p$. Here, again, we stress that the fact that all $\lambda$ are equal to one does not matter as long as alpha == NULL since we scale the vector automatically. Note that this option is only here for academic interest and to highlight the fact that SLOPE is a generalization of the lasso. There are more efficient packages, such as glmnet and biglasso, for fitting the lasso.

Solvers

There are currently two solvers available for SLOPE: FISTA (Beck and Teboulle 2009) and ADMM (Boyd et al. 2008). FISTA is available for families but ADMM is currently only available for family = "gaussian".

References


See Also

plot.SLOPE(), plotDiagnostics(), score(), predict.SLOPE(), trainSLOPE(), coef.SLOPE(), print.SLOPE(), print.SLOPE(), deviance.SLOPE(), sortedL1Prox()

Examples

# Gaussian response, default lambda sequence
fit <- SLOPE(bodyfat$x, bodyfat$y)

# Poisson response, OSCAR-type lambda sequence
fit <- SLOPE(
  abalone$x,
  abalone$y,
  family = "poisson",
  lambda = "oscar",
  theta1 = 1,
  theta2 = 0.9
)
# Multinomial response, custom alpha and lambda
m <- length(unique(wine$y)) - 1
p <- ncol(wine$x)

alpha <- 0.005
lambda <- exp(seq(log(2), log(1.8), length.out = p * m))

fit <- SLOPE(
  wine$x,
  wine$y,
  family = "multinomial",
  lambda = lambda,
  alpha = alpha
)

---

sortedL1Prox  

Sorted L1 Proximal Operator

Description

The proximal operator for the Sorted L1 Norm, which is the penalty function in SLOPE. It solves the problem

$$\arg\min_{x} \left( J(x, \lambda) + \frac{1}{2} ||x - v||_2^2 \right)$$

where $J(x, \lambda)$ is the Sorted L1 Norm.

Usage

sortedL1Prox(x, lambda, method = c("stack", "pava"))

Arguments

- **x** A vector. In SLOPE, this is the vector of coefficients.
- **lambda** A non-negative and decreasing sequence of weights for the Sorted L1 Norm. Needs to be the same length as x.
- **method** Method used in the prox. "stack" is a stack-based algorithm (Algorithm 4 in Bogdan et al.). "pava" is the PAVA algorithm used in isotonic regression (also Algorithm 3 in Bogdan et al.).

Value

An evaluation of the proximal operator at x and lambda.

Source

### Description

A data set of the attributes of 382 students in secondary education collected from two schools. The goal is to predict the grade in math and Portuguese at the end of the third period. See the cited sources for additional information.

### Usage

- student

### Format

382 observations from 13 variables represented as a list consisting of a binary factor response matrix \( y \) with two responses: Portuguese and math for the final scores in period three for the respective subjects. The list also contains \( x \): a sparse feature matrix of class ‘dgCMatrix’ with the following variables:

- **school_ms**: student’s primary school, 1 for Mousinho da Silveira and 0 for Gabriel Pereira
- **sex**: sex of student, 1 for male
- **age**: age of student
- **urban**: urban (1) or rural (0) home address
- **large_family**: whether the family size is larger than 3
- **cohabitation**: whether parents live together
- **Medu**: mother’s level of education (ordered)
- **Fedu**: father’s level of education (ordered)
- **Mjob_health**: whether the mother was employed in health care
- **Mjob_other**: whether the mother was employed as something other than the specified job roles
- **Mjob_services**: whether the mother was employed in the service sector
- **Mjob_teacher**: whether the mother was employed as a teacher
- **Fjob_health**: whether the father was employed in health care
- **Fjob_other**: whether the father was employed as something other than the specified job roles
- **Fjob_services**: whether the father was employed in the service sector
- **Fjob_teacher**: whether the father was employed as a teacher
- **reason_home**: school chosen for being close to home
- **reason_other**: school chosen for another reason
- **reason_rep**: school chosen for its reputation
- **nursery**: whether the student attended nursery school
- **internet**: Pwhether the student has internet access at home
Preprocessing

All of the grade-specific predictors were dropped from the data set. (Note that it is not clear from the source why some of these predictors are specific to each grade, such as which parent is the student’s guardian.) The categorical variables were dummy-coded. Only the final grades (G3) were kept as dependent variables, whilst the first and second period grades were dropped.

Source


See Also

Other datasets: abalone, bodyfat, heart, wine

---

**trainSLOPE**

*Train a SLOPE model*

**Description**

This function trains a model fit by SLOPE() by tuning its parameters through cross-validation.

**Usage**

```r
trainSLOPE(
  x,
  y,
  q = 0.2,
  number = 10,
  repeats = 1,
  measure = c("mse", "mae", "deviance", "misclass", "auc"),
  ...
)
```

**Arguments**

- `x` - the design matrix, which can be either a dense matrix of the standard *matrix* class, or a sparse matrix inheriting from *Matrix::sparseMatrix*. Data frames will be converted to matrices internally.
- `y` - the response, which for `family = "gaussian"` must be numeric; for `family = "binomial"` or `family = "multinomial"`, it can be a factor.
trainSLOPE

q  parameter controlling the shape of the lambda sequence, with usage varying depending on the type of path used and has no effect is a custom lambda sequence is used. Must be greater than 1e-6 and smaller than 1.

number  number of folds (cross-validation)
repeats  number of repeats for each fold (for repeated k-fold cross validation)
measure  measure to try to optimize; note that you may supply multiple values here and that, by default, all the possible measures for the given model will be used.
...  other arguments to pass on to SLOPE()

Details

Note that by default this method matches all of the available metrics for the given model family against those provided in the argument measure. Collecting these measures is not particularly demanding computationally so it is almost always best to leave this argument as it is and then choose which argument to focus on in the call to plot.TrainedSLOPE().

Value

An object of class "TrainedSLOPE", with the following slots:

summary  a summary of the results with means, standard errors, and 0.95 confidence levels
data  the raw data from the model training
optima  a data.frame of the best (mean) values for the different metrics and their corresponding parameter values
measure  a data.frame listing the used metrics and their labels
model  the model fit to the entire data set
call  the call

Parallel operation

This function uses the foreach package to enable parallel operation. To enable this, simply register a parallel backend using, for instance, doParallel::registerDoParallel() from the doParallel package before running this function.

See Also

foreach::foreach(), plot.TrainedSLOPE()

Other model-tuning: caretSLOPE(), plot.TrainedSLOPE()

Examples

# 8-fold cross-validation repeated 5 times
m <- trainSLOPE(subset(mtcars, select = c("mpg", "drat", "wt")),
mtcars$hp,
q = c(0.1, 0.2),
number = 8,
repeats = 5,
measure = "mse"
### Wine

**Wine cultivars**

<table>
<thead>
<tr>
<th>wine</th>
<th>Wine cultivars</th>
</tr>
</thead>
</table>

#### Description

A data set of results from chemical analysis of wines grown in Italy from three different cultivars.

#### Usage

wine

#### Format

178 observations from 13 variables represented as a list consisting of a categorical response vector $y$ with three levels: $A$, $B$, and $C$ representing different cultivars of wine as well as $x$: a sparse feature matrix of class ‘dgCMatrix’ with the following variables:

- **alcohol** alcoholic content
- **malic** malic acid
- **ash** ash
- **alcalinity** alcalinity of ash
- **magnesium** magnesium
- **phenols** total phenols
- **flavanoids** flavanoids
- **nonflavanoids** nonflavanoid phenols
- **proanthocyanins** proanthocyanins
- **color** color intensity
- **hue** hue
- **dilution** OD280/OD315 of diluted wines
- **proline** proline

#### Source


[https://raw.githubusercontent.com/hadley/rminds/master/1-data/wine.csv](https://raw.githubusercontent.com/hadley/rminds/master/1-data/wine.csv)

[https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/multiclass.html#wine](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/multiclass.html#wine)

#### See Also

Other datasets: abalone, bodyfat, heart, student
Index

* SLOPE-methods
  coef.SLOPE, 5
  deviance.SLOPE, 6
  plot.SLOPE, 8
  predict.SLOPE, 11
  print.SLOPE, 13
  score, 15

* datasets
  abalone, 2
  bodyfat, 3
  heart, 7
  student, 24
  wine, 27

* model-tuning
  caretSLOPE, 4
  plot.TrainedSLOPE, 9
  trainSLOPE, 25

abalone, 2, 4, 8, 25, 27
bodyfat, 3, 8, 25, 27
caret::train(), 4, 5
caretSLOPE, 4, 9, 26
coef.SLOPE, 5, 6, 8, 12, 13, 16
coef.SLOPE(), 12, 22
deviance.SLOPE, 5, 6, 8, 12, 13, 16
deviance.SLOPE(), 22
foreach::foreach(), 26
heart, 3, 4, 7, 25, 27
lattice::panel.xyplot(), 9
lattice::print.trellis(), 9
lattice::xyplot(), 8–10
Matrix::sparseMatrix, 17, 25
plot.SLOPE, 5, 6, 8, 12, 13, 16

plot.SLOPE(), 22
plot.TrainedSLOPE, 5, 9, 26
plot.TrainedSLOPE(), 26
plotDiagnostics, 10
plotDiagnostics(), 8, 22
predict.BinomialSLOPE (predict.SLOPE), 11
predict.GaussianSLOPE (predict.SLOPE), 11
predict.MultinomialSLOPE (predict.SLOPE), 11
predict.PoissonSLOPE (predict.SLOPE), 11
predict.SLOPE, 5, 6, 8, 11, 13, 16
predict.SLOPE(), 5, 16, 22
print(), 13
print.SLOPE, 5, 6, 8, 12, 13, 16
print.SLOPE(), 13, 22
print.TrainedSLOPE (print.SLOPE), 13

regularizationWeights, 13
score, 5, 6, 8, 12, 13, 15
score(), 22
SLOPE, 16
SLOPE(), 5, 6, 8, 10–16, 20, 25, 26
sortedL1Prox, 23
sortedL1Prox(), 18, 22
stats::predict(), 12
stats::predict.glm(), 12
stats::update(), 5, 12
student, 3, 4, 8, 24, 27

trainSLOPE, 5, 9, 25
trainSLOPE(), 4, 5, 9, 22

wine, 3, 4, 8, 25, 27