Package: savvySh (via r-universe)

October 21, 2026

Type Package

Title Slab and Shrinkage Linear Regression Estimation

Version 0.1.1

Description Implements a suite of shrinkage estimators for multivariate linear regression to improve estimation stability and predictive accuracy. Provides methods including the Stein estimator, Diagonal Shrinkage, the general Shrinkage estimator (solving a Sylvester equation), and Slab Regression (Simple and Generalized). These methods address Stein's paradox by introducing structured bias to reduce variance without requiring cross-validation, except for 'ShrinkageRR' where the intensity is chosen by minimizing an explicit Mean Squared Error (MSE) criterion. Methods are based on Asimit, V., Cidota, M. A., Chen, Z., and Asimit, J. (2025) <https://openaccess.city.ac.uk/id/eprint/35005/>.

License GPL (>= 3)

Encoding UTF-8

URL https://Ziwei-ChenChen.github.io/savvySh, https://github.com/Ziwei-ChenChen/savvySh

BugReports https://github.com/Ziwei-ChenChen/savvySh/issues

RoxygenNote 7.3.2

Depends R (>= 3.6.0)

Imports Matrix, glmnet, MASS, expm, mnormt, stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

Repository https://ziwei-chenchen.r-universe.dev

Date/Publication 2026-03-18 16:50:39 UTC

RemoteUrl https://github.com/ziwei-chenchen/savvysh

RemoteRef HEAD

RemoteSha 1c8076bce3971f2a3ade6496bcce7456db9c926d

Contents

coef.savvySh_model 2
cybersickness_10lags 3
cybersickness_raw 4
cybersickness_row 5
predict.savvySh_model 5
print.savvySh_model 7
returns_441 9
savvySh 9
summary.savvySh_model 12

Index **14**

`coef.savvySh_model` *Extract Coefficients for a Slab and Shrinkage Linear Regression Model*

Description

Extracts the regression coefficients from a `savvySh_model` object. You may specify one or more shrinkage estimators through the `estimator` parameter. If no estimator is specified, the function returns coefficients for all available estimators as a named list.

Usage

```
## S3 method for class 'savvySh_model'
coef(object, estimator = NULL, ...)
```

Arguments

`object`	A fitted `savvySh_model` object produced by `savvySh`.
`estimator`	A character vector naming one or more estimators from which to extract coefficients. Valid names are those stored in `object$coefficients` (e.g., `"St"`, `"DSh"`, `"Sh"`, `"SR"`, `"GSR"`, etc.). If `NULL`, coefficients for all available estimators are returned.
`...`	Additional arguments passed to `predict.savvySh_model`.

Details

This function internally calls `predict.savvySh_model` with `type = "coefficients"` to retrieve the desired coefficient estimates. If multiple estimators are requested (or if none is specified, in which case all are returned), the output is a named list in which each element is a numeric vector of coefficients. The coefficient vectors are named according to whether an intercept is present (for Linear shrinkage, no intercept). If a single estimator is specified, a single named numeric vector is returned.

Value

A named numeric vector of regression coefficients if a single estimator is specified, or a named list of such vectors if multiple estimators are requested.

Author(s)

Ziwei Chen, Vali Asimit, Marina Anca Cidota, Jennifer Asimit
Maintainer: Ziwei Chen <ziwei.chen.3@citystgeorges.ac.uk>

See Also

`predict.savvySh_model` for generating predictions, `savvySh` for fitting slab and shrinkage linear models.

Examples

```
# Generate simulated data for example
set.seed(123)
x <- matrix(rnorm(100 * 5), 100, 5)
y <- rnorm(100)

# Fit a Multiplicative shrinkage model
fit <- savvySh(x, y, model_class = "Multiplicative", include_Sh = TRUE)

# Extract coefficients for all available estimators (St and DSh by default)
all_coefs <- coef(fit)

# Extract coefficients for a specific estimator
st_coefs <- coef(fit, estimator = "St")
print(st_coefs)
```

---

cybersickness_10lags *Lagged Physiological Covariates for Cybersickness Prediction*

Description

This dataset is derived from 'cybersickness_row' by preprocessing the original physiological measurements. For each of the 22 physiological signals, 10 lagged features are created for each participant and time point. These lagged covariates, from time steps t-1 to t-10, are used as predictors for regression modeling in cybersickness studies.

Usage

```
cybersickness_10lags
```

Format

An object of class `data.frame` with 25663 rows and 132 columns.

Details

The preprocessing includes: 1. Creating lagged covariates: For each physiological signal Xi (for i = 2 to 23), new variables are created for values at previous time steps, including Xi(t-1), Xi(t-2), ..., Xi(t-10). 2. To avoid overlap between outcome and covariates, the last 10 rows for each participant are removed.

This preprocessing follows the steps outlined in https://github.com/shovonis/CyberSicknessClassification/tree/master/data_preprocessing.

A data frame with 25663 rows and 132 columns: - Intercept: Intercept column (all 1s). - X2(t-10) to X23(t-1): Lagged features for 22 physiological measurements, 10 lags per variable.

`cybersickness_raw` *Raw Physiological Measurements for Cybersickness Study*

Description

Raw data from a cybersickness study including heart rate, heart rate variability, and their percent changes.

Usage

```
cybersickness_raw
```

Format

An object of class `data.frame` with 25893 rows and 17 columns.

Details

A data frame with 25893 rows and 17 columns: - Epoch: Time interval index. - HR: Heart Rate. - PC_HR: Percent change in Heart Rate. - HR_MIN: Minimum heart rate in the epoch. - HR_MAX: Maximum heart rate in the epoch. - HRV: Heart Rate Variability. - PC_HRV: Percent change in HRV. - HRV_MIN: Minimum HRV in the epoch. - HRV_MAX: Maximum HRV in the epoch. - ...: Other columns include similar physiological signals.

The original data can be found at: https://github.com/shovonis/CyberSicknessClassification/blob/master/lstm_regression/data/raw_data.csv

cybersickness_row *Raw Physiological Measurements for Cybersickness Study*

Description

Raw data from a cybersickness study including heart rate, heart rate variability, and their percent changes.

Usage

```
cybersickness_row
```

Format

An object of class `data.frame` with 25893 rows and 17 columns.

Details

A data frame with 25893 rows and 17 columns: Epoch: Time interval index. HR: Heart Rate. PC_HR: Percent change in Heart Rate. HR_MIN: Minimum heart rate in the epoch. HR_MAX: Maximum heart rate in the epoch. HRV: Heart Rate Variability. PC_HRV: Percent change in HRV. HRV_MIN: Minimum HRV in the epoch. HRV_MAX: Maximum HRV in the epoch. ...: Other columns include similar physiological signals.

predict.savvySh_model *Predict Method for Slab and Shrinkage Linear Regression Models*

Description

Generate predictions (fitted values) or extract regression coefficients from a `savvySh_model` object returned by `savvySh`. This function allows you to specify one or more shrinkage estimators (via the `estimator` parameter) available in the model. If no estimator is specified, all available estimators are used and their results are returned in a named list.

Usage

```
## S3 method for class 'savvySh_model'
predict(
  object,
  newx = NULL,
  type = c("response", "coefficients"),
  estimator = NULL,
  ...
)
```

Arguments

`object`	A fitted `savvySh_model` object produced by `savvySh`.
`newx`	A numeric matrix of new predictor data for which to generate predictions. This argument is required if `type = "response"` and is ignored if `type = "coefficients"`.
`type`	A character string specifying the output type. Options are `"response"` to return predicted values and `"coefficients"` to extract regression coefficient vectors. Defaults to `"response"`.
`estimator`	A character vector naming one or more shrinkage estimator(s) to use. These must match names present in `object$coefficients`. If `NULL`, all available estimators are used.
`...`	Additional arguments (currently unused).

Details

The behavior depends on the value of `type`:

`"response"`**:** Generates predicted values using the coefficient estimates from the specified shrinkage estimator(s) for new data supplied via `newx`.

`"coefficients"`**:** Extracts the regression coefficient vector(s) corresponding to the specified estimator(s). Coefficient names are assigned based on whether an intercept is present (for Linear shrinkage, no intercept).

If no `estimator` is specified, the function returns results for all available estimators as a named list. If a single estimator is specified (or only one is provided in the vector), the result is returned as a numeric vector (for coefficients) or a numeric vector of predictions (for response).

Value

If `type = "response"`, the function returns:

- A numeric vector of predicted values if exactly one estimator is specified;
- Otherwise, a named list of numeric vectors, one for each specified estimator.

If `type = "coefficients"`, the function returns:

- A named numeric vector of regression coefficients if exactly one estimator is specified;
- Otherwise, a named list of numeric vectors corresponding to each specified estimator.

Author(s)

Ziwei Chen, Vali Asimit, Marina Anca Cidota, Jennifer Asimit
Maintainer: Ziwei Chen <ziwei.chen.3@citystgeorges.ac.uk>

See Also

`savvySh` for fitting slab and shrinkage linear models, `coef.savvySh_model` for direct coefficient extraction.

Examples

```
# Generate simulated data
set.seed(123)
x <- matrix(rnorm(100 * 5), 100, 5)
y <- rnorm(100)

# Fit a Multiplicative shrinkage model
fit <- savvySh(x, y, model_class = "Multiplicative")

# Generate predictions for new data
new_x <- matrix(rnorm(10 * 5), 10, 5)
preds <- predict(fit, newx = new_x, type = "response")

# Extract coefficients for specific estimators
coefs_st <- predict(fit, type = "coefficients", estimator = "St")
```

---

print.savvySh_model *Print a Slab and Shrinkage Model Summary*

Description

Displays a concise summary of a fitted `savvySh_model` object, including the original function call, the chosen model class, the number of non-zero coefficients per estimator, and the optimal `lambda` value (if applicable). Additionally, it prints the coefficients for the specified estimator(s) with user-specified precision.

Usage

```
## S3 method for class 'savvySh_model'
print(x, digits = max(3, getOption("digits") - 3), estimator = NULL, ...)
```

Arguments

`x`	A fitted `savvySh_model` object returned by `savvySh`.
`digits`	An integer specifying the number of significant digits to display when printing coefficients and `lambda`. Defaults to `max(3, getOption("digits") - 3)`.
`estimator`	A character vector naming one or more estimators for which coefficients should be printed. Valid names are those present in `x$coefficients` (e.g., `"St"`, `"DSh"`, `"Sh"`, `"SR"`, `"GSR"`, or `"ShrinkageRR"`). If `NULL`, coefficients for all estimators are printed.
`...`	Additional arguments passed to `print` (currently unused).

Details

This print method provides a quick diagnostic of the fitted model by showing:

Summary Metrics A table that includes, for each estimator, the number of non-zero coefficients and the optimal `lambda` (if applicable).

Coefficients For each selected estimator, the coefficients are printed with appropriate names: if an intercept is present, it is labeled `(Intercept)` and the remaining coefficients are labeled according to the predictor names.

If the user does not specify an estimator using the `estimator` argument, the function prints information for all available estimators stored in the model. If one or more estimators are specified, only those are printed, after verifying that they exist in `x$coefficients`.

The method invisibly returns a summary `data.frame` containing key metrics for each estimator.

Value

Invisibly returns a `data.frame` summarizing each selected estimator's name, number of non-zero coefficients, and the final `optimal_lambda` (if any).

Author(s)

Ziwei Chen, Vali Asimit, Marina Anca Cidota, Jennifer Asimit
Maintainer: Ziwei Chen <ziwei.chen.3@citystgeorges.ac.uk>

See Also

`savvySh` for fitting slab and shrinkage linear models, `coef.savvySh_model` and `predict.savvySh_model` for extracting coefficients and generating predictions.

Examples

```
# Generate simulated data for demonstration
set.seed(123)
x <- matrix(rnorm(100 * 5), 100, 5)
y <- rnorm(100)

# Fit a Multiplicative shrinkage model
fit <- savvySh(x, y, model_class = "Multiplicative", include_Sh = TRUE)

# Default print: shows summary metrics and coefficients for all estimators
print(fit)

# Print with specific digits and only for one estimator
print(fit, digits = 4, estimator = "St")
```

returns_441 *Stock Return Panel: 441 Stocks Across 6037 Dates*

Description

This dataset contains returns for 441 stocks (identified by their PERMNO codes), adjusted for dividends, over 6037 time periods. Each column corresponds to a unique stock, and the first column is the date of observation.

Usage

```
returns_441
```

Format

An object of class `data.frame` with 6037 rows and 442 columns.

Details

A data frame with 6037 rows and 442 columns:

Date: Date of the return observation (format: YYYY-MM-DD). PERMNO_*: Each remaining column corresponds to one stock's return.

savvySh *Slab and Shrinkage Linear Regression Estimation*

Description

This function estimates coefficients in a linear regression model using several shrinkage methods, including Multiplicative Shrinkage, Slab Regression, Linear shrinkage, and Shrinkage Ridge Regression. Each method gives estimators that balance bias and variance by applying shrinkage to the ordinary least squares (OLS) solution. The shrinkage estimators are computed based on different assumptions about the data.

Usage

```
savvySh(x, y, model_class = c("Multiplicative", "Slab", "Linear", "ShrinkageRR"),
               v = 1, lambda_vals = NULL, nlambda = 100, folds = 10,
               foldid = FALSE, include_Sh = FALSE, exclude = NULL)
```

Arguments

x	A matrix of predictor variables.
y	A vector of response variable.
model_class	A character string specifying the shrinkage model to use. Options can choose from "Multiplicative", "Slab", "Linear", and "ShrinkageRR". The default is "Multiplicative". If the user supplies more than one model, a warning is issued and only the first option is used.
v	A numeric value controlling the strength of shrinkage for the SR estimator in the "Slab" model. Must be a positive number. Default is 1.
lambda_vals	A vector of lambda values for RR. This is used only when multicollinearity (rank deficiency) is detected and "ShrinkageRR" is not selected. If NULL, a default sequence is used.
nlambda	The number of lambda values to use for cross-validation if lambda_vals is NULL. Only used when multicollinearity is present and "ShrinkageRR" is not called. The default is 100.
folds	Number of folds for cross-validation in RR. This is applicable only if multicollinearity occurs and "ShrinkageRR" is not chosen. The default is 10 and must be an integer >= 3.
foldid	Logical. If TRUE, saves the fold assignments in the output when multicollinearity is detected and "ShrinkageRR" is not used. The default is FALSE.
include_Sh	Logical. If TRUE, includes the Sh estimator in the "Multiplicative" model. The default is FALSE.
exclude	A vector specifying columns to exclude from the predictors. The default is NULL.

Details

The *Slab and Shrinkage Linear Regression Estimation* methodology provides four classes of shrinkage estimators that reduce variance in the OLS solution by introducing a small, structured bias. These methods handle overfitting, collinearity, and high-dimensional scenarios by controlling how and where the coefficients are shrunk. Each class offers a distinct strategy for controlling instability and improving mean squared error (MSE) in linear models, tailored for different modeling contexts specified in the model_class argument. Note that if the user provides more than one option in model_class, only the first option is used, and a warning is issued.

Model Classes:

Multiplicative Shrinkage: This class includes three estimators that use the OLS coefficients as a starting point and apply *multiplicative* adjustments:

- St - *Stein estimator*, which shrinks all coefficients toward zero by a single global factor. This aims to reduce MSE while keeping the overall bias fairly uniform across coefficients.
- DSh - *Diagonal Shrinkage*, assigning an individual factor to each coefficient based on its variance. This yields more targeted shrinkage than the global approach and often achieves a lower MSE.
- Sh - *Shrinkage estimator* that solves a Sylvester equation for a full (non-diagonal) shrinkage matrix. It is more flexible but also more computationally demanding. Included only if include_Sh = TRUE.

Slab Regression: *Slab Regression* applies an adaptive quadratic penalty term to the OLS objective:

- SR - *Simple Slab Regression*, which modifies the OLS objective by adding a penalty in a fixed direction (often the constant vector). This penalty is controlled by v and does not require cross-validation. It can be viewed as a special case of the generalized lasso but focuses on smooth (quadratic) rather than ℓ_1 regularization.
- GSR - *Generalized Slab Regression*, extending SR by allowing shrinkage along multiple directions. Typically, these directions correspond to the eigenvectors of the design covariance matrix, effectively shrinking principal components.

Linear Shrinkage: The *Linear Shrinkage (LSh)* estimator forms a convex combination of the OLS estimator (through the origin) and a target estimator that assumes uncorrelated predictors (*diagonal* approximation of the covariance). This approach is simpler than a full matrix method and is well-suited for standardized data where the intercept is not needed.

Shrinkage Ridge Regression: The *Shrinkage Ridge Regression (SRR)* extends standard RR by shrinking the design covariance matrix toward a spherical target (i.e., a diagonal matrix with equal entries). This additional regularization stabilizes the eigenvalues and yields more robust coefficient estimates, particularly when the predictors lie close to a low-dimensional subspace.

Value

A list containing the following elements:

- `call` The matched function call.
- `model` The data frame of y and x used in the analysis.
- `optimal_lambda` If x is full rank, this value is 0. If x is rank-deficient, it is the chosen RR lambda from cross-validation.
- `model_class` The selected model class.
- `coefficients` A list of estimated coefficients for each applicable estimator in the model_class.
- `fitted_values` A list of fitted values for each estimator.
- `pred_MSE` A list of prediction MSEs for each estimator.
- `ridge_results (optional)` A list containing detailed results from RR, used when multicollinearity (rank deficiency) is detected in x and the "ShrinkageRR" is not called. This element is included only when RR is applied instead of OLS due to the rank deficiency of x. It contains:
  - `lambda_range` The range of lambda values used in the RR cross-validation.
  - `cvm` A vector of cross-validated MSEs for each lambda in lambda_range.
  - `cvsd` The standard deviation of the cross-validated MSEs for each lambda.
  - `ridge_coefficients` A matrix of coefficients from RR at each lambda value, with each column representing the coefficients corresponding to a specific lambda.

Author(s)

Ziwei Chen, Vali Asimit, Marina Anca Cidota, Jennifer Asimit
Maintainer: Ziwei Chen <ziwei.chen.3@citystgeorges.ac.uk>

References

Asimit, V., Cidota, M. A., Chen, Z., & Asimit, J. (2025). *Slab and Shrinkage Linear Regression Estimation*. Retrieved from https://openaccess.city.ac.uk/id/eprint/35005/

Examples

```
# 1. Simple Multiplicative Shrinkage example
set.seed(123)
x <- matrix(rnorm(100 * 5), 100, 5)
y <- rnorm(100)
fit_mult <- savvySh(x, y, model_class = "Multiplicative", include_Sh = TRUE)
print(fit_mult)

# 2. Slab Regression example
fit_slab <- savvySh(x, y, model_class = "Slab", v = 2)
coef(fit_slab, estimator = "GSR")

# 3. Linear Shrinkage (standardized data recommended)
x_centered <- scale(x, center = TRUE, scale = FALSE)
y_centered <- scale(y, center = TRUE, scale = FALSE)
fit_linear <- savvySh(x_centered, y_centered, model_class = "Linear")

# 4. Shrinkage Ridge Regression
fit_srr <- savvySh(x, y, model_class = "ShrinkageRR")
predict(fit_srr, newx = matrix(rnorm(10 * 5), 10, 5), type = "response")
```

---

summary.savvySh_model *Summarize a Slab and Shrinkage Linear Regression Model*

Description

Provides a comprehensive summary for one or more shrinkage estimators contained within a `savvySh_model` object produced by `savvySh`. The summary includes estimated coefficients, confidence intervals, residual statistics, R-squared measures, F-statistics, and information criteria (AIC, BIC) for each specified estimator.

Usage

```
## S3 method for class 'savvySh_model'
summary(object, estimator = NULL, ...)
```

Arguments

`object`	A fitted model object of class `savvySh_model`, produced by `savvySh`.
`estimator`	A character vector naming one or more estimators to summarize (e.g., `"St"`, `"DSh"`, `"SR"`, `"GSR"`, `"Sh"`, etc.). If `NULL` (default), summaries for all available estimators are printed.
`...`	Additional arguments (currently unused).

Details

For each estimator present in the `savvySh_model` object (or for the user-specified subset), this function computes:

- A summary of the residual distribution (quantiles).
- A coefficient table including estimates, standard errors, t-values, p-values, and confidence intervals.
- Residual standard error and degrees of freedom.
- R-squared and adjusted R-squared measures.
- F-statistic (and its p-value) for testing overall regression significance.
- Information criteria (AIC, BIC) and deviance for model fit.

These results are printed in sequence for the selected estimator(s). If no estimator is specified, summaries for all available estimators are printed.

Value

Invisibly returns a `data.frame` summarizing key metrics for each estimator (including estimator name, number of non-zero coefficients, and optimal `lambda` if available).

Author(s)

Ziwei Chen, Vali Asimit, Marina Anca Cidota, Jennifer Asimit
Maintainer: Ziwei Chen <ziwei.chen.3@citystgeorges.ac.uk>

See Also

`savvySh` for fitting slab and shrinkage linear models, `predict.savvySh_model` for generating predictions, `coef.savvySh_model` for extracting coefficients directly.

Examples

```
# Generate simulated data for demonstration
set.seed(123)
x <- matrix(rnorm(100 * 5), 100, 5)
y <- rnorm(100)

# Fit a Slab Regression model
fit <- savvySh(x, y, model_class = "Slab")

# Print a detailed summary for all estimators (SR and GSR)
summary(fit)

# Summarize only a specific estimator
summary(fit, estimator = "GSR")
```

Index

∗ **datasets**
    cybersickness_10lags, 3
    cybersickness_raw, 4
    cybersickness_row, 5
    returns_441, 9

coef.savvySh_model, 2, *6*, *8*, *13*
cybersickness_10lags, 3
cybersickness_raw, 4
cybersickness_row, 5

predict.savvySh_model, *2*, *3*, 5, *8*, *13*
print, *7*
print.savvySh_model, 7

returns_441, 9

savvySh, *3*, *6*, *8*, 9, *13*
summary.savvySh_model, 12