Package: psvmSDR (via r-universe)

October 14, 2026

Type Package

Title Unified Principal Sufficient Dimension Reduction Package

Version 3.0.1

Maintainer Jungmin Shin <c16267@gmail.com>

Description A unified and user-friendly framework for applying the principal sufficient dimension reduction methods for both linear and nonlinear cases. The package has an extendable power by varying loss functions for the support vector machine, even for an user-defined arbitrary function, unless those are convex and differentiable everywhere over the support (Li et al. (2011) <doi:10.1214/11-AOS932>). Also, it provides a real-time sufficient dimension reduction update procedure using the principal least squares support vector machine (Artemiou et al. (2021) <doi:10.1016/j.patcog.2020.107768>).

License GPL-2

Encoding UTF-8

Imports stats, graphics

Suggests testthat (>= 3.0.0)

RoxygenNote 7.3.3

Config/testthat/edition 3

NeedsCompilation no

Author Jungmin Shin [aut, cre], Seung Jun Shin [aut], Andreas Artemiou [aut]

Repository https://c16267.r-universe.dev

Date/Publication 2026-02-16 17:20:02 UTC

RemoteUrl https://github.com/cran/psvmSDR

RemoteRef HEAD

RemoteSha e3fabaafdb2b012f2946525c7b529b0126ff26ca

Contents

npsdr 2
npsdr_x 4
plot.npsdr 5
plot.psdr 7
psdr 8
psdr_bic 11
rtpsdr 13

Index **15**

`npsdr` *A unified Principal sufficient dimension reduction method via kernel trick*

Description

This function extends principal SDR to nonlinear relationships between predictors and the response using a kernel feature map. The kernel basis is constructed internally using a data-driven number of basis functions, and the working matrix is formed analogously to linear principal SDR but in the transformed feature space.

Users may choose from built-in loss functions or provide a custom loss through the same interface as psdr(). The method supports both continuous and binary responses and can visualize the nonlinear sufficient predictors.

The output contains the kernel basis object, the working matrix M, eigenvalues and eigenvectors, and detailed fitting metadata.

Usage

```
npsdr(
  x,
  y,
  loss = "svm",
  h = 10,
  lambda = 1,
  b = floor(length(y)/3),
  eps = 1e-05,
  max.iter = 100,
  eta = 0.1,
  mtype = "m",
  plot = TRUE
)
```

Arguments

x	data matrix
y	either continuous or (+1,-1) typed binary response vector
loss	pre-specified loss functions belongs to "svm", "logit", "l2svm", "wsvm", "qr", "asls", "wlogit", "wl2svm", "lssvm", "wlssvm", and user-defined loss function object also can be used formed by inside double (or single) quotation mark. Default is 'svm'.
h	unified control for slicing or weighting; accepts either an integer or a numeric vector.
lambda	hyperparameter for the loss function. default value is 1
b	number of basis functions for a kernel trick, floor(length(y)/3) is default
eps	threshold for stopping iteration with respect to the magnitude of derivative, default value is 1.0e-4
max.iter	maximum iteration number for the optimization process. default value is 30
eta	learning rate for gradient descent method. default value is 0.1
mtype	type of margin, either "m" or "r" refer margin and residual, respectively (See, Table 1 in the pacakge manuscript). When one use user-defined loss function this argument should be specified. Default is "m".
plot	If TRUE then it produces scatter plots of Y versus the first sufficient predictor. The default is FALSE.

Value

An object of class "npsdr" containing:

- x, y: input data
- M: working matrix
- evalues, evectors: eigen-decomposition of M
- obj.psi: kernel basis object from get.psi()
- fit: metadata (loss, h, lambda, eps, max.iter, eta, b, response.type, cutpoints, weight_cutpoints)

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

References

Artemiou, A. and Dong, Y. (2016) *Sufficient dimension reduction via principal lq support vector machine, Electronic Journal of Statistics 10: 783–805*.
Artemiou, A., Dong, Y. and Shin, S. J. (2021) *Real-time sufficient dimension reduction through principal least squares support vector machines, Pattern Recognition 112: 107768*.
Kim, B. and Shin, S. J. (2019) *Principal weighted logistic regression for sufficient dimension reduction in binary classification, Journal of the Korean Statistical Society 48(2): 194–206*.
Li, B., Artemiou, A. and Li, L. (2011) *Principal support vector machines for linear and nonlinear*

sufficient dimension reduction, Annals of Statistics 39(6): 3182–3210.
Soale, A.-N. and Dong, Y. (2022) *On sufficient dimension reduction via principal asymmetric least squares, Journal of Nonparametric Statistics 34(1): 77–94.*
Wang, C., Shin, S. J. and Wu, Y. (2018) *Principal quantile regression for sufficient dimension reduction with heteroscedasticity, Electronic Journal of Statistics 12(2): 2114–2140.*
Shin, S. J., Wu, Y., Zhang, H. H. and Liu, Y. (2017) *Principal weighted support vector machines for sufficient dimension reduction in binary classification, Biometrika 104(1): 67–81.*
Li, L. (2007) *Sparse sufficient dimension reduction, Biometrika 94(3): 603–613.*

See Also

npsdr_x, psdr, rtpsdr

Examples

```
set.seed(1)
n <- 200;
p <- 5;
x <- matrix(rnorm(n*p, 0, 2), n, p)
y <- 0.5*sqrt((x[,1]^2+x[,2]^2))*(log(x[,1]^2+x[,2]^2))+ 0.2*rnorm(n)
obj_kernel <- npsdr(x, y, plot=FALSE)
print(obj_kernel)
summary(obj_kernel)
plot(obj_kernel)
```

---

npsdr_x *Reconstruct estimated sufficient predictors for new data*

Description

Computes the nonlinear sufficient predictors $\hat{\phi}(\mathbf{x})$ for a new data matrix using a previously fitted npsdr object.

This function evaluates the learned kernel-based sufficient dimension reduction (SDR) mapping on new observations. Given a fitted nonlinear SDR model $\hat{\phi}$ estimated from npsdr(), the function computes:

$$\hat{Z} = \hat{\phi}(X_{\text{new}}) = \Psi(X_{\text{new}})^\top \hat{V}_{1:d},$$

where $\Psi(\cdot)$ is the kernel feature map constructed from the training data, and $\hat{V}_{1:d}$ contains the first d eigenvectors of the estimated working matrix M. These eigenvectors span the estimated central subspace in the kernel-transformed space.

This enables users to extract sufficient predictors for downstream tasks such as visualization, classification, regression, or clustering on new data, without re-estimating the SDR model.

Usage

```
npsdr_x(object, newdata, d = 2)
```

Arguments

object	The object from function npsdr
newdata	new data $\mathbf{X}$
d	structural dimensionality. d=2 is default.

Value

the value of the estimated nonlinear mapping $\phi(\cdot)$ is applied to newdata X with dimension d is returned.

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

See Also

npsdr

Examples

```
set.seed(1)
n <- 200; n.new <- 300
p <- 5;
h <- 20;
x <- matrix(rnorm(n*p, 0, 2), n, p)
y <- 0.5*sqrt((x[,1]^2+x[,2]^2))*(log(x[,1]^2+x[,2]^2))+ 0.2*rnorm(n)
new.x <- matrix(rnorm(n.new*p, 0, 2), n.new, p)
obj_kernel <- npsdr(x, y)
z_new <- npsdr_x(object=obj_kernel, newdata=new.x)
dim(z_new)
```

---

plot.npsdr *Plot sufficient predictors from an npsdr object*

Description

Creates diagnostic scatter plots of nonlinear sufficient predictors produced by npsdr(). The function visualizes the estimated transformed directions and optionally overlays a lowess smoothing curve for continuous responses.

Additional graphical arguments can be provided. These plots help assess nonlinear structure in the data and evaluate how effectively the kernel SDR method reduces dimensionality.

Usage

```
## S3 method for class 'npsdr'
plot(
  x,
  ...,
  d = 1,
  lowess = TRUE,
  col = NULL,
  line.col = "red",
  pch = 16,
  lwd = 1.2,
  xlab = NULL,
  ylab = NULL
)
```

Arguments

x	object from npsdr()
...	additional graphical parameters for plot()
d	number of sufficient predictors to plot (default = 1)
lowess	draw a lowess curve for continuous responses (default = TRUE)
col	point color(s)
line.col	color for lowess smoothing line (default = "red")
pch	point character (default = 16)
lwd	line width for smoothing (default = 1.2)
xlab	x-axis label (default depends on predictor index)
ylab	y-axis label (default = "Y" for continuous)

Value

A scatter plot with sufficient predictors.

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

See Also

npsdr_x, npsdr

Examples

```
set.seed(1)
n <- 200;
p <- 5;
x <- matrix(rnorm(n*p, 0, 2), n, p)
```

```
y <-  x[,1]/(0.5 + (x[,2] + 1)^2) + 0.2*rnorm(n)
obj_kernel <- npsdr(x, y, plot=FALSE)
plot(obj_kernel, d = 1)
```

`plot.psdr` *Plot sufficient predictors from a psdr object*

Description

Produces scatter plots of the sufficient predictors obtained from psdr(). For continuous responses, the function plots Y versus each selected sufficient predictor along with an optional lowess curve. For binary responses, a two-dimensional scatter plot of the first two sufficient predictors is produced with class-specific point colors.

Additional graphical parameters may be passed to the underlying plot() function. The plot is intended as a diagnostic tool to visualize the estimated central subspace and assess how well the sufficient predictors capture the relationship between X and Y.

Usage

```
## S3 method for class 'psdr'
plot(
  x,
  ...,
  d = 1,
  lowess = TRUE,
  col = NULL,
  line.col = "red",
  pch = 16,
  lwd = 1.2,
  xlab = NULL,
  ylab = NULL
)
```

Arguments

`x`	object from the function `psdr()`
`...`	Additional graphical parameters passed to `plot()`.
`d`	number of sufficient predictors. Default is 1.
`lowess`	draw a locally weighted scatterplot smoothing curve. Default is TRUE.
`col`	color vector for points (optional; defaults depend on response type)
`line.col`	color for lowess smoothing line (default = "red")
`pch`	plotting character (default = 16)
`lwd`	line width for smoothing curve (default = 1.2)
`xlab`	label for x-axis (default depends on d)
`ylab`	label for y-axis (default depends on response type)

Value

A scatter plot with sufficient predictors and the lowess curve is overlayed as default.

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

See Also

psdr_bic, psdr

Examples

```
set.seed(1)
n <- 200; p <- 5;
x <- matrix(rnorm(n*p, 0, 2), n, p)
y <-  x[,1]/(0.5 + (x[,2] + 1)^2) + 0.2*rnorm(n)
obj <- psdr(x, y)
plot(obj)
```

---

psdr *Unified linear principal sufficient dimension reduction methods*

Description

This function implements a unified framework for linear principal SDR methods. It provides a single interface that covers many existing principal-machine approaches, such as principal SVM, weighted SVM, logistic, quantile, and asymmetric least squares SDR. The method estimates the central subspace by constructing a working matrix M derived from user-specified loss functions, slicing or weighting schemes, and regularization.

The function is designed for both continuous responses and binary classification (with any two-level coding). Users may choose among several built-in loss functions or supply a custom loss function. Two examples of the usage of user-defined losses are presented below (u represents a margin):

`mylogit <- function(u, ...) log(1+exp(-u))`,

`myls <- function(u ...) u^2`.

Argument `u` is a function variable (any character is possible) and the argument `mtype` for `psdr()` determines a type of a margin, either (`type="m"`) or (`type="r"`) method. `type="m"` is a default. Users have to change `type="r"`, when applying residual type loss. Any additional parameters of the loss can be specified via `...` argument.

The output includes the estimated eigenvalues and eigenvectors of M, which form the basis of the estimated central subspace, as well as detailed metadata used to summarize model fitting and diagnostics.

Usage

```
psdr(
  x,
  y,
  loss = "svm",
  h = 10,
  lambda = 1,
  eps = 1e-05,
  max.iter = 100,
  eta = 0.1,
  mtype = "m",
  plot = FALSE
)
```

Arguments

x	input matrix, of dimension nobs x nvars; each row is an observation vector.
y	response variable, either continuous or binary (any 2-level coding; e.g., -1/1, 0/1, 1/2, TRUE/FALSE, factor/character).
loss	pre-specified loss functions belongs to "svm", "logit", "l2svm", "wsvm", "qr", "asls", "wlogit", "wl2svm", "lssvm", "wlssvm", and user-defined loss function object also can be used formed by inside double (or single) quotation mark. Default is 'svm'.
h	unified control for slicing or weighting; accepts either an integer or a numeric vector.
lambda	regularization parameter (default 1).
eps	convergence threshold on parameter change (default 1e-5).
max.iter	maximum number of iterations (default 100).
eta	learning rate for gradient descent (default 0.1).
mtype	a margin type, which is either margin ("m") or residual ("r") (See, Table 1 in the manuscript). Only need when user-defined loss is used. Default is "m".
plot	logical; if TRUE, produces diagnostic plot.

Value

An object of S3 class "psdr" containing

- M: working matrix
- evalues, evectors: eigen decomposition of M
- fit: metadata (n, p, ytype, hyperparameters, per-slice iteration/convergence info)

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

References

Artemiou, A. and Dong, Y. (2016) *Sufficient dimension reduction via principal lq support vector machine, Electronic Journal of Statistics 10: 783–805*.
Artemiou, A., Dong, Y. and Shin, S. J. (2021) *Real-time sufficient dimension reduction through principal least squares support vector machines, Pattern Recognition 112: 107768*.
Kim, B. and Shin, S. J. (2019) *Principal weighted logistic regression for sufficient dimension reduction in binary classification, Journal of the Korean Statistical Society 48(2): 194–206*.
Li, B., Artemiou, A. and Li, L. (2011) *Principal support vector machines for linear and nonlinear sufficient dimension reduction, Annals of Statistics 39(6): 3182–3210*.
Soale, A.-N. and Dong, Y. (2022) *On sufficient dimension reduction via principal asymmetric least squares, Journal of Nonparametric Statistics 34(1): 77–94*.
Wang, C., Shin, S. J. and Wu, Y. (2018) *Principal quantile regression for sufficient dimension reduction with heteroscedasticity, Electronic Journal of Statistics 12(2): 2114–2140*.
Shin, S. J., Wu, Y., Zhang, H. H. and Liu, Y. (2017) *Principal weighted support vector machines for sufficient dimension reduction in binary classification, Biometrika 104(1): 67–81*.
Li, L. (2007) *Sparse sufficient dimension reduction, Biometrika 94(3): 603–613*.

See Also

psdr_bic, rtpsdr

Examples

```
## ----------------------------
## Linear PM
## ----------------------------
set.seed(1)
n <- 200; p <- 5;
x <- matrix(rnorm(n*p, 0, 2), n, p)
y <-  x[,1]/(0.5 + (x[,2] + 1)^2) + 0.2*rnorm(n)
y.tilde <- sign(y)
obj <- psdr(x, y)
print(obj)
plot(obj, d=2)

## --------------------------
## User defined cutoff points
## --------------------------
obj_cut <- psdr(x, y, h = c(0.1, 0.3, 0.5, 0.7))
print(obj_cut)

## --------------------------------
## Linear PM (Binary classification)
## --------------------------------
obj_wsvm <- psdr(x, y.tilde, loss="wsvm")
plot(obj_wsvm)

## ----------------------------
## User-defined loss function
## ----------------------------
mylogistic <- function(u) log(1+exp(-u))
```

```
psdr(x, y, loss="mylogistic")

## ----------------------------
## Real-data example: iris (binary subset)
## ----------------------------
iris_binary <- droplevels(subset(iris, Species %in% c("setosa", "versicolor")))
psdr(x = iris_binary[, 1:4], y = iris_binary$Species, plot = TRUE)
```

`psdr_bic` *Structural dimension selection for principal SDR*

Description

This function selects the structural dimension d of a fitted psdr model using the BIC-type criterion proposed by Li, Artemiou and Li (2011). The criterion evaluates cumulative eigenvalues of the working matrix, applying a penalty that depends on the tuning parameter rho and the sample size.

Selects the structural dimension d of a principal SDR model using the BIC-type criterion of Li et al. (2011):

$$G(d) = \sum_{j=1}^{d} v_j \; - \; \rho \frac{d \log n}{\sqrt{n}} \, v_1,$$

where v_j are the eigenvalues of the working matrix M.

To improve robustness, cross-validation is used to choose ρ based on the stability of the selected structural dimension across folds. Specifically, for each candidate ρ, the data are split into K folds, and a dimension estimate $\hat{d}^{(k)}(\rho)$ is obtained from fold k.

The CV stability metric is defined as

$$\mathrm{Var}_{CV}(\rho) = \frac{1}{K} \sum_{k=1}^{K} \left\{ \hat{d}^{(k)}(\rho) - \overline{d}(\rho) \right\}^2,$$

where

$$\overline{d}(\rho) = \frac{1}{K} \sum_{k=1}^{K} \hat{d}^{(k)}(\rho).$$

The value of ρ that minimizes $\mathrm{Var}_{CV}(\rho)$ is selected, yielding a dimension estimate that is both theoretically justified (via the BIC-type criterion) and empirically stable (via cross-validation).

The function returns the selected ρ, the corresponding estimated dimension d, the matrix of BIC-type criterion values, and the CV-based stability metrics.

Usage

```
psdr_bic(
  obj,
  rho_grid = seq(0.001, 0.05, length = 10),
  cv_folds = 5,
  plot = TRUE,
  seed = 123,
  ...
)
```

Arguments

obj	A fitted psdr object.
rho_grid	Numeric vector of candidate ρ values. Default seq(0.001, 0.05, length=10).
cv_folds	Number of cross-validation folds for stability evaluation. Default is 5.
plot	Logical; if TRUE, plots the BIC-type criterion curve and CV stability.
seed	Random seed for reproducibility.
...	Additional graphical arguments for plot.

Value

A list of class "psdr_bic" containing:

- rho_star - selected rho minimizing cross-validated variation
- d_hat - estimated structural dimension
- G_values - matrix of BIC-type scores for each rho
- cv_variation - variation (variance) of d_hat across folds
- fold_dhat - per-fold estimated dimensions

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

References

Li, B., Artemiou, A. and Li, L. (2011) *Principal support vector machines for linear and nonlinear sufficient dimension reduction, Annals of Statistics 39(6): 3182–3210.*

See Also

psdr

Examples

```
set.seed(1)
n <- 200; p <- 5;
x <- matrix(rnorm(n*p), n, p)
y <- x[,1]/(0.5+(x[,2]+1)^2)+0.2*rnorm(n)
fit <- psdr(x, y, loss="svm")
bic_out <- psdr_bic(fit, rho_grid=seq(0.05, 0.1, length=5), cv_folds=5)
bic_out$d_hat
```

`rtpsdr` *Real time sufficient dimension reduction through principal least squares SVM*

Description

This function implements a real-time version of principal SDR based on least squares SVM loss. It is intended for streaming or sequential data settings where new observations arrive continuously and re-fitting the full SDR model would be computationally expensive.

After an initial psdr or rtpsdr fit is obtained, this function updates the working matrix M, slice statistics, and eigen-decomposition efficiently using only the new batch of data. The method supports both regression and binary classification, automatically choosing the appropriate LS-SVM variant.

The returned object includes cumulative sample size, updated mean vector, slice coefficients, intermediate matrices required for updates, and the resulting central subspace basis.

Usage

```
rtpsdr(x, y, obj = NULL, h = 10, lambda = 1)
```

Arguments

`x`	x in new data
`y`	y in new data, y is continuous
`obj`	the latest output object from the `rtpsdr`
`h`	unified control for slicing or weighting; accepts either an integer or a numeric vector.
`lambda`	hyperparameter for the loss function. default is set to 1.

Value

An object of class `c("rtpsdr","psdr")` containing:

- `x`, `y`: latest batch data
- `M`: working matrix
- `evalues`, `evectors`: eigen-decomposition of `M` (central subspace basis)

- `N`: cumulative sample size
- `Xbar`: cumulative mean vector
- `r`: slice-specific coefficient matrix
- `A`: new A part for update. See Artemiou et. al., (2021)
- `loss`: "lssvm" (continuous) or "wlssvm" (binary)
- `fit`: metadata (mode="realtime", H, cutpoints, weight_cutpoints, lambda, etc.)

Author(s)

Jungmin Shin, <c16267@gmail.com>, Seung Jun Shin, <sjshin@korea.ac.kr>, Andreas Artemiou <artemiou@uol.ac.cy>

References

Artemiou, A. and Dong, Y. (2016) *Sufficient dimension reduction via principal lq support vector machine, Electronic Journal of Statistics 10: 783–805*.
Artemiou, A., Dong, Y. and Shin, S. J. (2021) *Real-time sufficient dimension reduction through principal least squares support vector machines, Pattern Recognition 112: 107768*.
Kim, B. and Shin, S. J. (2019) *Principal weighted logistic regression for sufficient dimension reduction in binary classification, Journal of the Korean Statistical Society 48(2): 194–206*.
Li, B., Artemiou, A. and Li, L. (2011) *Principal support vector machines for linear and nonlinear sufficient dimension reduction, Annals of Statistics 39(6): 3182–3210*.
Soale, A.-N. and Dong, Y. (2022) *On sufficient dimension reduction via principal asymmetric least squares, Journal of Nonparametric Statistics 34(1): 77–94*.
Wang, C., Shin, S. J. and Wu, Y. (2018) *Principal quantile regression for sufficient dimension reduction with heteroscedasticity, Electronic Journal of Statistics 12(2): 2114–2140*.
Shin, S. J., Wu, Y., Zhang, H. H. and Liu, Y. (2017) *Principal weighted support vector machines for sufficient dimension reduction in binary classification, Biometrika 104(1): 67–81*.
Li, L. (2007) *Sparse sufficient dimension reduction, Biometrika 94(3): 603–613*.

See Also

`psdr`, `npsdr`

Examples

```
set.seed(1)
p <- 5; m <- 300; B <- 3
obj <- NULL
for (b in 1:B) {
  x <- matrix(rnorm(m*p), m, p)
  y <- x[,1]/(0.5+(x[,2]+1)^2) + 0.2*rnorm(m)
  obj <- rtpsdr(x, y, obj=obj, h=8, lambda=1)
}
print(obj)
summary(obj)
```

Index

npsdr, 2, *5*, *6*, *14*
npsdr_x, 4, *4*, *6*

plot.npsdr, 5
plot.psdr, 7
psdr, *4*, 8, *8*, *12*, *14*
psdr_bic, *8*, *10*, 11

rtpsdr, *4*, *10*, 13