Package ‘spcr’

September 28, 2021

Type Package

Title Sparse Principal Component Regression

Version 2.1

Date 2021-09-28

Author Shuichi Kawano

Maintainer Shuichi Kawano <`skawano@ai.lab.uec.ac.jp`>

Description The sparse principal component regression is computed. The regularization parameters are optimized by cross-validation.

License GPL (>= 2)

URL `https://sites.google.com/site/shuichikawanoen/`

Repository CRAN

NeedsCompilation yes

Date/Publication 2021-09-28 11:30:02 UTC

R topics documented:

`cv.spcr` *Cross-validation for spcr*

Description

This function performs cross-validation for spcr. `cv.spcr` enables us to determine two regularization parameters λ_β and λ_γ objectively.

Usage

```
cv.spcr(x, y, k, w=0.1, xi=0.01, nfolds=5, adaptive=FALSE,
center=TRUE, scale=FALSE, lambda.B.length=10, lambda.gamma.length=10,
lambda.B=NULL, lambda.gamma=NULL)
```

Arguments

`x`	A data matrix.
`y`	A response vector.
`k`	The number of principal components.
`w`	Weight parameter with $0 \leq w \leq 1$. The default is 0.1.
`xi`	The elastic net mixing parameter with $0 \leq \alpha \leq 1$. The default is 0.01.
`nfolds`	The number of folds. The default is 5.
`adaptive`	If `"TRUE"`, the adaptive SPCR is used.
`center`	If `"TRUE"`, the data matrix is centered.
`scale`	If `"TRUE"`, the data matrix is scaled.
`lambda.B.length`	The number of candidates for the parameter λ_β. The default is 10.
`lambda.gamma.length`	The number of candidates for the parameter λ_γ. The default is 10.
`lambda.B`	Optional user-supplied candidates for the parameter λ_β. The default is NULL.
`lambda.gamma`	Optional user-supplied candidates for the parameter λ_γ. The default is NULL.

Value

`lambda.gamma.seq`	The values of `lambda.gamma` in the fit.
`lambda.B.seq`	The values of `lambda.B` in the fit.
`CV.mat`	Matrix of the mean values of cross-validation. The row shows a sequence of `lambda.gamma`. The column shows a sequence of `lambda.B`.
`lambda.gamma.cv`	The value of `lambda.gamma` selected by cross-validation.
`lambda.B.cv`	The value of `lambda.B` selected by cross-validation.
`cvm`	The minimum of the mean cross-validated error.

Author(s)

Shuichi Kawano
`<skawano@ai.lab.uec.ac.jp>`

References

Kawano, S., Fujisawa, H., Takada, T. and Shiroishi, T. (2015). *Sparse principal component regression with adaptive loading. Compuational Statistics & Data Analysis, 89, 192–203.*

See Also

`spcr`

Examples

```
#data
n <- 50
np <- 5
set.seed(1)
nu0 <- c(-1, 1)
x <- matrix( rnorm(np*n), n, np )
e <- rnorm(n)
y <- nu0[1]*x[ ,1] + nu0[2]*x[ ,2] + e

#fit
cv.spcr.fit <- cv.spcr(x=x, y=y, k=2)
cv.spcr.fit

#fit (adaptive SPCR)
cv.adaspcr.fit <- cv.spcr(x=x, y=y, k=2, adaptive=TRUE)
cv.adaspcr.fit
```

`cv.spcrglm` *Cross-validation for spcr-glm*

Description

This function performs cross-validation for SPCR-glm. `cv.spcrglm` enables us to determine two regularization parameters λ_β and λ_γ objectively.

Usage

```
cv.spcrglm(x, y, k, family=c("binomial","poisson","multinomial"),
w=0.1, xi=0.01, nfolds=5, adaptive=FALSE, q=1, center=TRUE,
scale=FALSE, lambda.B.length=10, lambda.gamma.length=10,
lambda.B=NULL, lambda.gamma=NULL)
```

Arguments

`x`	A data matrix.
`y`	A response vector.
`k`	The number of principal components.
`family`	Response type.
`w`	Weight parameter with $w \geq 0$. The default is 0.1.
`xi`	The elastic net mixing parameter with $0 \leq \alpha \leq 1$. The default is 0.01.
`nfolds`	The number of folds. The default is 5.

`adaptive`	If `"TRUE"`, the adaptive SPCR-glm (aSPCR-glm) is used.
`q`	The tuning parameter that controls weights in aSPCR-glm. The default is 1.
`center`	If `"TRUE"`, the data matrix is centered.
`scale`	If `"TRUE"`, the data matrix is scaled.
`lambda.B.length`	The number of candidates for the parameter λ_β. The default is 10.
`lambda.gamma.length`	The number of candidates for the parameter λ_γ. The default is 10.
`lambda.B`	Optional user-supplied candidates for the parameter λ_β. The default is NULL.
`lambda.gamma`	Optional user-supplied candidates for the parameter λ_γ. The default is NULL.

Value

`lambda.gamma.seq`	The values of `lambda.gamma` in the fit.
`lambda.B.seq`	The values of `lambda.B` in the fit.
`CV.mat`	Matrix of the mean values of cross-validation. The row shows a sequence of `lambda.gamma`. The column shows a sequence of `lambda.B`.
`lambda.gamma.cv`	The value of `lambda.gamma` selected by cross-validation.
`lambda.B.cv`	The value of `lambda.B` selected by cross-validation.
`cvm`	The minimum of the mean cross-validated error.

Author(s)

Shuichi Kawano
`<skawano@ai.lab.uec.ac.jp>`

References

Kawano, S., Fujisawa, H., Takada, T. and Shiroishi, T. (2018). *Sparse principal component regression for generalized linear models. Compuational Statistics & Data Analysis, 124, 180–196.*

See Also

`spcrglm`

Examples

```
# binomial
n <- 100
np <- 3
nu0 <- c(-1, 1)
set.seed(4)
x <- matrix( rnorm(np*n), n, np )
y <- rbinom(n,1,1-1/(1+exp(  (nu0[1]*x[ ,1] + nu0[2]*x[ ,2]  ))))
cv.spcrglm.fit <- cv.spcrglm(x=x, y=y, k=1, family="binomial")
```

```
cv.spcrglm.fit

# Poisson
set.seed(5)
y <- rpois(n, 1)
cv.spcrglm.fit <- cv.spcrglm(x=x, y=y, k=1, family="poisson")
cv.spcrglm.fit

# multinomial
set.seed(4)
y <- sample(1:4, n, replace=TRUE)
cv.spcrglm.fit <- cv.spcrglm(x=x, y=y, k=1, family="multinomial")
cv.spcrglm.fit
```

`spcr` *Fit a sparse principal component regression (SPCR)*

Description

This function computes a principal component regression model via sparse regularization.

Usage

```
spcr(x, y, k, lambda.B, lambda.gamma, w=0.1, xi=0.01,
adaptive=FALSE, center=TRUE, scale=FALSE)
```

Arguments

`x`	A data matrix.
`y`	A response vector.
`k`	The number of principal components.
`lambda.B`	The regularization parameter for the parameter B.
`lambda.gamma`	The regularization parameter for the coefficient vector γ.
`w`	Weight parameter with $0 \leq w \leq 1$. The default is 0.1.
`xi`	The elastic net mixing parameter with $0 \leq \alpha \leq 1$. The default is 0.01.
`adaptive`	If `"TRUE"`, the adaptive SPCR is used.
`center`	If `"TRUE"`, the data matrix is centered.
`scale`	If `"TRUE"`, the data matrix is scaled.

Value

`loadings.B`	the loading matrix B
`gamma`	the coefficient
`gamma0`	intercept
`loadings.A`	the loading matrix A

Author(s)

Shuichi Kawano
`<skawano@ai.lab.uec.ac.jp>`

References

Kawano, S., Fujisawa, H., Takada, T. and Shiroishi, T. (2015). *Sparse principal component regression with adaptive loading. Compuational Statistics & Data Analysis, 89, 192–203.*

See Also

`cv.spcr`

Examples

```
#data
n <- 100
np <- 5
set.seed(4)
nu0 <- c(-1, 1)
x <- matrix( rnorm(np*n), n, np )
e <- rnorm(n)
y <- nu0[1]*x[ ,1] + nu0[2]*x[ ,2] + e

#fit
spcr.fit <- spcr(x=x, y=y, k=2, lambda.B=6, lambda.gamma=2)
spcr.fit

#fit (adaptive SPCR)
adaspcr.fit <- spcr(x=x, y=y, k=2, lambda.B=6, lambda.gamma=2, adaptive=TRUE)
adaspcr.fit
```

`spcrglm` *Fit a sparse principal component regression for generalized linear models (SPCR-glm)*

Description

This function computes a principal component regression for generalized linear models via sparse regularization.

Usage

```
spcrglm(x, y, k, family=c("binomial","poisson","multinomial"), lambda.B,
lambda.gamma, w=0.1, xi=0.01, adaptive=FALSE, q=1, center=TRUE, scale=FALSE)
```

Arguments

`x`	A data matrix.
`y`	A response data.
`k`	The number of principal components.
`family`	Response type.
`lambda.B`	The regularization parameter for the parameter B.
`lambda.gamma`	The regularization parameter for the coefficient vector γ.
`w`	Weight parameter with $w \geq 0$. The default is 0.1.
`xi`	The elastic net mixing parameter with $0 \leq \alpha \leq 1$. The default is 0.01.
`adaptive`	If `"TRUE"`, the adaptive SPCR-glm (aSPCR-glm) is used.
`q`	The tuning parameter that controls weights in aSPCR-glm. The default is 1.
`center`	If `"TRUE"`, the data matrix is centered.
`scale`	If `"TRUE"`, the data matrix is scaled.

Value

`loadings.B`	the loading matrix B
`gamma`	the coefficient
`gamma0`	intercept
`loadings.A`	the loading matrix A

Author(s)

Shuichi Kawano
`<skawano@ai.lab.uec.ac.jp>`

References

Kawano, S., Fujisawa, H., Takada, T. and Shiroishi, T. (2018). *Sparse principal component regression for generalized linear models. Compuational Statistics & Data Analysis, 124, 180–196.*

See Also

`cv.spcrglm`

Examples

```
# binomial
n <- 100
np <- 5
nu0 <- c(-1, 1)
set.seed(4)
x <- matrix( rnorm(np*n), n, np )
y <- rbinom(n,1,1-1/(1+exp(  (nu0[1]*x[ ,1] + nu0[2]*x[ ,2]  ))))
spcrglm.fit <- spcrglm(x=x, y=y, k=2, family="binomial", lambda.B=2, lambda.gamma=1)
```

```
spcrglm.fit

# Poisson
set.seed(4)
y <- rpois(n, exp(  (nu0[1]*x[ ,1] + nu0[2]*x[ ,2] ) ))
spcrglm.fit <- spcrglm(x=x, y=y, k=2, family="poisson", lambda.B=2, lambda.gamma=1)
spcrglm.fit

# multinomial
set.seed(4)
y <- sample(1:4, n, replace=TRUE)
spcrglm.fit <- spcrglm(x=x, y=y, k=2, family="multinomial", lambda.B=2, lambda.gamma=2)
spcrglm.fit
```

Index