Package 'MVOPR'

March 31, 2025

Type Package

Title Multi-View Orthogonal Projection Regression for Multi-Modality Integration

Version 2.0.0

Description
Implements the 'MVOPR' (Multi-View Orthogonal Projection Regression) method for robust variable selection and integration of multi-modality data.

License GPL-2 | GPL-3

Encoding UTF-8

Imports ncvreg, rrpack

RoxygenNote 7.3.2

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

URL https://arxiv.org/abs/2503.16807

NeedsCompilation no

Author Zongrui Dai [aut, cre] (<https://orcid.org/0000-0002-7893-5004>), Gen Li [aut]

Maintainer Zongrui Dai <daizr@umich.edu>

Repository CRAN

Date/Publication 2025-03-31 17:30:10 UTC

Contents

MVOPR2 *Multi-View Orthogonal Projection Regression for two modalities*

Description

Fit Multi-View Orthogonal Projection Regression for two modalities with Lasso, MCP, SCAD. The function is capable for linear, logistic, and poisson regression.

Usage

```
MVOPR2(
  M1,
  M2,
  Y,
  RRR_Control = list(Sparsity = TRUE, nrank = 10, ic.type = "GIC"),
  family = "gaussian",
  penalty = "lasso"
)
```

Arguments

M1	A numeric matrix (n x p) for the first modality.
M2	A numeric matrix (n x q) for the second modality. Assumes 'M2' is correlated to 'M1' via a low-rank matrix.
Y	A numeric response vector of length 'n', connected to 'M1' and 'M2'.
RRR_Control	A list to control the fitting for reduced rank regression.
	Sparsity Logical. If 'TRUE', performs Sparse Orthogonal Factor Regression (SOFAR); otherwise, a reduced-rank regression model is fitted.
	nrank Integer. Maximum rank to be searched for the reduced-rank model.
	ic.type Character. Model selection criterion: '"AIC"', '"BIC"', or '"GIC"'.
family	Either "gaussian", "binomial", or "poisson", depending on the response.
penalty	The penalty to be applied in the outcome model Y to M1 and M2. Either "MCP" (the default), "SCAD", or "lasso".

Value

A list containing:

fitY Results for Outcome regression (Y~M1+M2). A fitted object from 'cv.ncvreg', which contains the penalized regression results for 'Y'.

fitM2 Results for reduced-rank regression (M2~M1).The fitted reduced-rank regression model from 'rrpack'.

CoefY A vector of estimated regression coefficients for 'M1' and 'M2' on 'Y'.

coefM2 A matrix of estimated regression coefficients for 'M1' on 'M2'.

`rank` An integer indicates the estimated rank of the reduced-rank regression.

`P` A projection matrix used to extract the orthogonal components of 'M1'.

`M1s` Transformed version of 'M1' after projection.

`M2s` Transformed version of 'M2' after removing the effect of 'M1'.

References

Dai, Z., Huang, Y. J., & Li, G. (2025). Multi-View Orthogonal Projection Regression with Application in Multi-omics Integration. arXiv preprint arXiv:2503.16807. Available at <https://arxiv.org/abs/2503.16807>

Examples

```
## Simulation.1
p = 100; q = 100; n = 200
rank = 3

beta = c(rep(c(rep(1,5),rep(0,95)),2))
M1 = matrix(rnorm(p*n),n,p)

U = matrix(rnorm(rank*p),p,rank)
V = matrix(rnorm(rank*q),rank,q)
B = U %*% V
E = matrix(rnorm(q*n),n,q)
M2 = M1 %*% B + E
Y = cbind(M1,M2) %*% matrix(beta,p+q,1)

Fit = MVOPR2(M1,M2,Y,RRR_Control = list(Sparsity = FALSE))

## Result for variable selection
print(data.frame(Truecoef = beta,estimate = Fit$CoefY[2:(p+q+1)]))

## Plot the pathway and cv error in outcome model
oldpar <- par(mfrow = c(1, 2))
on.exit(par(oldpar))
plot(Fit$fitY$fit)
plot(Fit$fitY)
```

`MVOPR3` *Multi-View Orthogonal Projection Regression for three modalities*

Description

Fit Multi-View Orthogonal Projection Regression for three modalities with Lasso, MCP, SCAD. The function is capable for linear, logistic, and poisson regression.

Usage

```
MVOPR3(
  M1,
  M2,
  M3,
  Y,
  RRR_Control = list(Sparsity = TRUE, nrank = 10, ic.type = "GIC"),
  family = "gaussian",
  penalty = "lasso"
)
```

Arguments

M1	A numeric matrix (n x p1) for the first modality.
M2	A numeric matrix (n x p2) for the second modality. Assumes 'M2' is correlated to 'M1' via a low-rank matrix.
M3	A numeric matrix (n x p3) for the third modality. Assumes 'M3' is correlated to 'M1' and 'M2' via a low-rank matrix.
Y	A numeric response vector of length 'n', connected to 'M1', 'M2', and 'M3'.
RRR_Control	A list to control the fitting for reduced rank regression.
	Sparsity Logical. If 'TRUE', performs Sparse Orthogonal Factor Regression (SOFAR); otherwise, a reduced-rank regression model is fitted.
	nrank Integer. Maximum rank to be searched for the reduced-rank model.
	ic.type Character. Model selection criterion: '"AIC"', '"BIC"', or '"GIC"'.
family	Either "gaussian", "binomial", or "poisson", depending on the response.
penalty	The penalty to be applied in the outcome model Y to M1 and M2. Either "MCP" (the default), "SCAD", or "lasso".

Value

A list containing:

fitY A fitted object from 'cv.ncvreg', containing the penalized regression results for 'Y'.

fitM2 The fitted reduced-rank regression ('sofar' or 'rrr' object) for 'M2' given 'M1'.

fitM3 The fitted reduced-rank regression ('sofar' or 'rrr' object) for 'M3' given 'M1' and 'M2'.

CoefY A vector of estimated regression coefficients for 'Y'.

coefM2 A matrix of estimated regression coefficients for 'M2' given 'M1'.

coefM3 A matrix of estimated regression coefficients for 'M3' given 'M1' and 'M2'.

rank1 An integer indicating the estimated rank of the reduced-rank regression for 'M2'.

rank2 An integer indicating the estimated rank of the reduced-rank regression for 'M3'.

P1 A projection matrix used to extract the orthogonal components of 'M1'.

P2 A projection matrix used to extract the orthogonal components of 'E2', which is the error term in the regression for 'M2' given 'M1'.

M1s A transformed version of ‘M1‘ after projection.

M2s A transformed version of ‘M2‘ after removing the effect of ‘M1‘ and projecting to the orthogonal space.

M3s A transformed version of ‘M3‘ after removing the effects of ‘M1‘ and ‘M2‘.

#’ @references Dai, Z., Huang, Y. J., & Li, G. (2025). Multi-View Orthogonal Projection Regression with Application in Multi-omics Integration. arXiv preprint arXiv:2503.16807. Available at <https://arxiv.org/abs/2503.16807>

Examples

```
## Simulation: three modalities
p1 = 50; p2 = 50; p3 = 50; n = 200
rank = 2

beta = c(rep(c(rep(1,5),rep(0,45)),3))
M1 = matrix(rnorm(p1*n),n,p1)

U1 = matrix(rnorm(rank*p1),p1,rank)
V1 = matrix(runif(rank*p2,-0.1,0.1),rank,p2)
B1 = U1 %*% V1

U2 = matrix(rnorm(rank*p1),p1,rank)
V2 = matrix(runif(rank*p2,-0.1,0.1),rank,p3)
B2 = U2 %*% V2

U3 = matrix(rnorm(rank*p2),p2,rank)
V3 = matrix(runif(rank*p2,-0.1,0.1),rank,p3)
B3 = U3 %*% V3

E1 = matrix(rnorm(p2*n),n,p2)
E2 = matrix(rnorm(p3*n),n,p3)

M2 = M1 %*% B1 + E1
M3 = M1 %*% B2 + M2 %*% B3 + E2
Y = cbind(M1,M2,M3) %*% matrix(beta,p1+p2+p3,1)

## Fit MVOPR with Lasso
Fit1 = MVOPR3(M1,M2,M3,Y,RRR_Control = list(Sparsity = FALSE),penalty = 'lasso')

## Fit MVOPR with MCP
Fit2 = MVOPR3(M1,M2,M3,Y,RRR_Control = list(Sparsity = FALSE),penalty = 'MCP')

## Fit MVOPR with SCAD
Fit3 = MVOPR3(M1,M2,M3,Y,RRR_Control = list(Sparsity = FALSE),penalty = 'SCAD')

## Compare the variable selection between Lasso, MCP, SCAD
print(data.frame(Lasso = Fit1$CoefY[2:151],MCP = Fit2$CoefY[2:151],SCAD = Fit3$CoefY[2:151],beta))
```

Index