Package ‘mixedLSR’

November 4, 2022

Title  Mixed, Low-Rank, and Sparse Multivariate Regression on High-Dimensional Data

Version  0.1.0

Description  Mixed, low-rank, and sparse multivariate regression (‘mixedLSR’) provides tools for performing mixture regression when the coefficient matrix is low-rank and sparse. ‘mixedLSR’ allows subgroup identification by alternating optimization with simulated annealing to encourage global optimum convergence. This method is data-adaptive, automatically performing parameter selection to identify low-rank substructures in the coefficient matrix.

License  MIT + file LICENSE

Encoding  UTF-8

RoxygenNote  7.2.1

Depends  R (>= 4.1.0)

Imports  grpreg, purrr, MASS, stats, ggplot2

Suggests  knitr, rmarkdown, mclust

VignetteBuilder  knitr

BugReports  https://github.com/alexanderjwhite/mixedLSR

URL  https://alexanderjwhite.github.io/mixedLSR/

NeedsCompilation  no

Author  Alexander White [aut, cre] (https://orcid.org/0000-0002-9117-1475), Sha Cao [aut] (https://orcid.org/0000-0002-8645-848X), Yi Zhao [ctb] (https://orcid.org/0000-0003-4766-5934), Chi Zhang [ctb] (https://orcid.org/0000-0001-9553-0925)

Maintainer  Alexander White <whitealj@iu.edu>

Repository  CRAN

Date/Publication  2022-11-04 20:00:02 UTC
### R topics documented:

- `bic_lsr` .......................... 2
- `mixed_lsr` .......................... 3
- `plot_lsr` .......................... 4
- `simulate_lsr` .......................... 5

## Index

```
bic_lsr .......................... 2
mixed_lsr .......................... 3
plot_lsr .......................... 4
simulate_lsr .......................... 5
```

---

### `bic_lsr`

**Compute Bayesian information criterion for a mixedLSR model**

### Description

Compute Bayesian information criterion for a mixedLSR model

### Usage

```r
bic_lsr(a, n, llik)
```

### Arguments

- `a`  
  A list of coefficient matrices.
- `n`  
  The sample size.
- `llik`  
  The log-likelihood of the model.

### Value

The BIC.

### Examples

```r
n <- 50
simulate <- simulate_lsr(n)
model <- mixed_lsr(simulate$x, simulate$y, k = 2, init_lambda = c(1,1), alt_iter = 0)
bic_lsr(model$A, n = n, model$llik)
```
Description

Mixed Low-Rank and Sparse Multivariate Regression for High-Dimensional Data

Usage

```r
mixed_lsr(
  x,
  y,
  k,
  nstart = 1,
  init_assign = NULL,
  init_lambda = NULL,
  alt_iter = 5,
  anneal_iter = 1000,
  em_iter = 1000,
  temp = 1000,
  mu = 0.95,
  eps = 1e-06,
  accept_prob = 0.95,
  sim_N = 200,
  verbose = TRUE
)
```

Arguments

- **x**: A matrix of predictors.
- **y**: A matrix of responses.
- **k**: The number of groups.
- **nstart**: The number of random initializations, the result with the maximum likelihood is returned.
- **init_assign**: A vector of initial assignments, NULL by default.
- **init_lambda**: A vector with the values to initialize the penalization parameter for each group, e.g., c(1,1,1). Set to NULL by default.
- **alt_iter**: The maximum number of times to alternate between the classification expectation maximization algorithm and the simulated annealing algorithm.
- **anneal_iter**: The maximum number of simulated annealing iterations.
- **em_iter**: The maximum number of EM iterations.
- **temp**: The initial simulated annealing temperature, temp > 0.
**plot_lsr**

mu

The simulated annealing decrease temperature fraction. Once the best configuration cannot be improved, reduce the temperature to \((\mu)T, 0 < \mu < 1\).

eps

The final simulated annealing temperature, \(\text{eps} > 0\).

accept_prob

The simulated annealing probability of accepting a new assignment \(0 < \text{accept_prob} < 1\). When closer to 1, trial assignments will only be small perturbation of the current assignment. When closer to 0, trial assignments are closer to random.

sim_N

The simulated annealing number of iterations for reaching equilibrium.

verbose

A boolean indicating whether to print to screen.

Value

A list containing the likelihood, the partition, the coefficient matrices, and the BIC.

Examples

```r
simulate <- simulate_lsr(50)
mixed_lsr(simulate$x, simulate$y, k = 2, init_lambda = c(1,1), alt_iter = 0)
```

---

**plot_lsr**

Heatmap Plot of the mixedLSR Coefficient Matrices

Description

Heatmap Plot of the mixedLSR Coefficient Matrices

Usage

`plot_lsr(a, abs = TRUE)`

Arguments

- `a`: A coefficient matrix from mixed_lsr model.
- `abs`: A boolean for taking the absolute value of the coefficient matrix.

Value

A ggplot2 heatmap of the coefficient matrix, separated by subgroup.

Examples

```r
simulate <- simulate_lsr()
plot_lsr(simulate$a)
```
Simulate Heterogeneous, Low-Rank, and Sparse Data

Usage

simulate_lsr(
  N = 100,
  k = 2,
  p = 30,
  m = 35,
  b = 1,
  d = 20,
  h = 0.2,
  case = "independent"
)

Arguments

N The sample size, default = 100.
k The number of groups, default = 2.
p The number of predictor features, default = 30.
m The number of response features, default = 35.
b The signal-to-noise ratio, default = 1.
d The singular value, default = 20.
h The lower bound for the singular matrix simulation, default = 0.2.
case The covariance case, "independent" or "dependent", default = "independent".

Value

A list of simulation values, including x matrix, y matrix, coefficients and true clustering assignments.

Examples

simulate_lsr()
Index

bic_lsr, 2
mixed_lsr, 3
plot_lsr, 4
simulate_lsr, 5