Package 'MVSKmod'

July 21, 2025

Type Package

Title Matrix-Variate Skew Linear Regression Models

Version 0.1.0

Maintainer Samuel Soon <samksoon2@gmail.com>

Description An implementation of the alternating expectation conditional maximization (AECM) algorithm for matrix-variate variance gamma (MVVG) and normal-inverse Gaussian (MVNIG) linear models. These models are designed for settings of multivariate analysis with clustered non-uniform observations and correlated responses. The package includes fitting and prediction functions for both models, and an example dataset from a periodontal on Gullah-speaking African Americans, with responses in gaad_res, and covariates in gaad_cov. For more details on the matrix-variate distributions used, see Gallaugher & McNicholas (2019) <doi:10.1016/j.spl.2018.08.012>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.3.1

Imports Bessel, clusterGeneration, DistributionUtils, matlib, maxLik, truncnorm, pracma

URL https://github.com/soonsk-vcu/MVSKmod

BugReports https://github.com/soonsk-vcu/MVSKmod/issues

NeedsCompilation no

Author Samuel Soon [aut, cre], Dipankar Bandyopadhyay [aut], Qingyang Liu [aut]

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2025-05-09 14:10:13 UTC

Contents

gaad_cov 2
gaad_res 3
gaad_theta_mvvg 3
MVNIGmod 4
MVVGmod 6
predict 8
theta_mvnig 9
theta_mvvg 9
X 10
Y 10

Index **11**

---

gaad_cov *GAAD Data*

Description

These data sets describe periodontal measurements performed on members of the Gullah-Speaking African American community.

Usage

```
gaad_cov
```

Format

Each is a list of matrices, with rows denoting tooth sites and columns denoting CAL/PPD response.

Details

`gaad_res` and `gaad_cov` contain the response and covariate matrices of the GAAD data.

Examples

```
gaad_cov
gaad_res
```

gaad_res *GAAD Data*

Description

These data sets describe periodontal measurements performed on members of the Gullah-Speaking African American community.

Usage

```
gaad_res
```

Format

Each is a list of matrices, with rows denoting tooth sites and columns denoting CAL/PPD response.

Details

`gaad_res` and `gaad_cov` contain the response and covariate matrices of the GAAD data.

Examples

```
gaad_cov
gaad_res
```

gaad_theta_mvvg *MVVG Parameter Format*

Description

This is an example of the format of input parameter list theta.

Usage

```
gaad_theta_mvvg
```

Format

List of model parameters.

Examples

```
gaad_theta_mvvg
```

`MVNIGmod` *AECM Estimation for Matrix-Variate Normal-Inverse Gaussian Models*

Description

This function fits MVNIG linear models for matrix-variate skew data with non-uniform data rows between subjects. Exchangeable observation row correlation and skewness structures are imposed to accommodate the varying row counts across matrices. Note that multiple restarts may be needed to account for unstable local maxima.

Usage

```
MVNIGmod(Y, X, theta_g = NULL, stopping = 0.001, max_iter = 50)
```

Arguments

`Y`	List of $n_i \times p$ response matrices. Matrices must have same number of columns.
`X`	List of $n_i \times q$ design matrices. Matrices must have same number of columns.
`theta_g`	List of parameters to pass as initial values in the AECM algorithm. If NULL, will be randomly generated. See Details for an in-depth explanation.
`stopping`	Stopping threshold for the L-infinity norm of differences in consecutive parameter space, evaluated at iteration $t+1$ as $\|\hat{\theta}^{t+1} - \hat{\theta}^{t}\|_{\infty}$. Default is 0.001
`max_iter`	Maximum number of iterations, default is 50.

Details

Fits the matrix-variate skew regression model

$$Y_i = X_i\Theta + E_i,$$

where each response Y_i is a $n_i \times p$ matrix that indexes n_i observations and p response variables. X_i corresponds to a $n_i \times q$ design matrix, and Θ corresponds to a $q \times p$ coefficient matrix. E_i corresponds to a $n_i \times p$ error matrix, following a matrix-variate variance-gamma distribution.

The model estimates MVVG parameters $\Theta, \underline{a}, r, \Psi, \tilde{\gamma}$ using the alternating expectation conditional maximization (AECM) algorithm, using the density

$$f(Y_i|M_i,\underline{a},r,\Psi,\tilde{\gamma},n_i,p) = \frac{2\exp[matlib :: tr(\Sigma_i^{-1}(Y_i-M_i)\Psi^{-1}A_i^T)+\tilde{\gamma}]}{(2\pi)^{\frac{n_ip}{2}+1}|\Sigma_i|^{\frac{p}{2}}|\Psi|^{\frac{n_i}{2}}}\left(\frac{\delta(Y_i;M_i,\Sigma_i,\Psi)+1}{\rho(A_i,\Sigma_i,\Psi)+\tilde{\gamma}^2}\right)^{-\frac{(1+n_ip)}{4}} \times K_{-\frac{(1+n_ip)}{2}}$$

where $A_i = \underline{1}_{n_i} \times \underline{a}^T$, $\Sigma_i = I_{n_i} + r(\underline{1}_{n_i}\underline{1}_{n_i}^T - I_{n_i})$, $\delta(X;M,\Sigma,\Psi) = matlib :: tr(\Sigma^{-1}(X-M)\Psi^{-1}(X-M)^T)$, $\rho(A,\Sigma,\Psi) = matlib :: tr(\Sigma^{-1}A\Psi^{-1}A^T)$, and $K_\nu(x)$ is the modified Bessel function of the second kind.

The structure of `theta_g` and parameter estimates returned by the function must be in the form of a list with the following named elements:

`Theta`: $q \times p$ coefficient matrix

`a`: $p \times 1$ skewness vector

`rho`: Compound symmetry parameter for row correlation matrix

`Psi`: $p \times p$ column covariance matrix

`tgamma`: Univariate mixing parameter

Value

MVNIGmod returns a list with the following elements:

`Iteration`: Number of iterations taken to convergence. `Inf` if convergence not reached.

`Starting Value`: List of initial parameter values.

`Final Value`: List of final parameter estimates.

`Stopping Criteria`: Vector of $|\hat{\theta}^{t+1} - \hat{\theta}^{t}|_{\infty}$ at each iteration.

`AIC`: Model AIC

`BIC`: Model BIC

Author(s)

Samuel Soon

Dipankar Bandyopadhyay

Qingyang Liu

Examples

```
MVNIGmod(Y,X,theta_mvnig)

set.seed(1234)
# num response variables
p <- ncol(gaad_res[[1]])
# num covariates
q <- ncol(gaad_cov[[1]])
# generate initial value to input, then run AECM with MVVG distribution
initial_mvnig_theta <- list(Theta = matrix(stats::rnorm(p*q), nrow = q, ncol = p),
                      A = rep(1,p),
                     rho = 0.3,
                     Psi = diag(p),
                     tgamma = 3)
MVNIGmod(gaad_res[1:30], gaad_cov[1:30], initial_mvnig_theta)
```

MVVGmod *AECM Estimation for Matrix-Variate Variance Gamma (MVVG) Models*

Description

This function fits MVVG linear models for matrix-variate skew data with non-uniform data rows between subjects. Exchangeable observation row correlation and skewness structures are imposed to accommodate the varying row counts across matrices. Note that multiple restarts may be needed to account for unstable local maxima.

Usage

```
MVVGmod(Y, X, theta_g = NULL, stopping = 0.001, max_iter = 50)
```

Arguments

Y	List of $n_i \times p$ response matrices. Matrices must have same number of columns.
X	List of $n_i \times q$ design matrices. Matrices must have same number of columns.
theta_g	List of parameters to pass as initial values in the AECM algorithm. If NULL, will be randomly generated. See Details for an in-depth explanation.
stopping	Stopping threshold for the L-infinity norm of differences in consecutive parameter space, evaluated at iteration $t+1$ as $\|\hat{\theta}^{t+1} - \hat{\theta}^{t}\|_{\infty}$. Default is 0.001
max_iter	Maximum number of iterations, default is 50.

Details

Fits the matrix-variate skew regression model

$$Y_i = X_i\Theta + E_i,$$

where each response Y_i is a $n_i \times p$ matrix that indexes n_i observations and p response variables. X_i corresponds to a $n_i \times q$ design matrix, and Θ corresponds to a $q \times p$ coefficient matrix. E_i corresponds to a $n_i \times p$ error matrix, following a matrix-variate variance-gamma distribution.

The model estimates MVVG parameters $\Theta, \underline{a}, r, \Psi, \gamma$ using the alternating expectation conditional maximization (AECM) algorithm, using the density

$$f(Y_i|X_i\Theta, \underline{a}, r, \Psi, \gamma, n_i, p) = \frac{2\gamma^{\gamma}\exp[matlib :: tr(\Sigma_i^{-1}(Y_i - X_i\Theta)\Psi^{-1}A_i^T)]}{(2\pi)^{n_ip/2}|\Sigma_i|^{p/2}|\Psi|^{n_i/2}\Gamma(\gamma)}\left(\frac{\delta(Y_i; X_i\Theta, \Sigma_i, \Psi)}{\rho(A_i, \Sigma_i, \Psi) + 2\gamma}\right)^{(\gamma - n_ip/2)/2} \times K_{(\gamma - n_ip}$$

where $A_i = \underline{1}_{n_i} \times \underline{a}^T$, $\Sigma_i = I_{n_i} + r(\underline{1}_{n_i}\underline{1}_{n_i}^T - I_{n_i})$, $\delta(X; M, \Sigma, \Psi) = matlib :: tr(\Sigma^{-1}(X - M)\Psi^{-1}(X - M)^T)$, $\rho(A, \Sigma, \Psi) = matlib :: tr(\Sigma^{-1}A\Psi^{-1}A^T)$, and $K_{\nu}(x)$ is the modified Bessel function of the second kind.

The structure of theta_g and parameter estimates returned by the function must be in the form of a list with the following named elements:

`Theta`: $q \times p$ coefficient matrix

`a`: $p \times 1$ skewness vector

`rho`: Compound symmetry parameter for row correlation matrix

`Psi`: $p \times p$ column covariance matrix

`gamma`: Univariate mixing parameter

Value

MVVGmod returns a list with the following elements:

`Iteration`: Number of iterations taken to convergence. `Inf` if convergence not reached.

`Starting Value`: List of initial parameter values.

`Final Value`: List of final parameter estimates.

`Stopping Criteria`: Vector of $|\hat{\theta}^{t+1} - \hat{\theta}^{t}|_{\infty}$ at each iteration.

`AIC`: Model AIC

`BIC`: Model BIC

Author(s)

Samuel Soon

Dipankar Bandyopadhyay

Qingyang Liu

Examples

```
MVVGmod(Y,X,theta_mvvg)

set.seed(1234)
# num response variables
p <- ncol(gaad_res[[1]])
# num covariates
q <- ncol(gaad_cov[[1]])
# generate initial value to input, then run AECM with MVVG distribution
initial_gaad_theta_mvvg <- list(Theta = matrix(stats::rnorm(p*q), nrow = q, ncol = p),
                      A = rep(1,p),
                     rho = 0.3,
                     Psi = diag(p),
                     gamma = 4)
MVVGmod(gaad_res[1:50], gaad_cov[1:50], initial_gaad_theta_mvvg)
```

predict *MVSK Model Prediction*

Description

Predicts response values given a list of covariate matrices and a model output from either MVVGmod or MVNIGmod.

Usage

```
predict(mod, X)
```

Arguments

mod	object outputted by either MVVGmod or MVNIGmod
X	Inputted covariate matrix

Value

Returns a list of predicted response matrices

Author(s)

Samuel Soon

Dipankar Bandyopadhyay

Qingyang Liu

Examples

```
set.seed(1234)
# num response variables
p <- ncol(gaad_res[[1]])
# num covariates
q <- ncol(gaad_cov[[1]])
# generate initial value to input, then run AECM with MVVG distribution
initial_mvnig_theta <- list(Theta = matrix(stats::rnorm(p*q), nrow = q, ncol = p),
                      A = rep(1,p),
                     rho = 0.3,
                     Psi = diag(p),
                     tgamma = 4)
mvnig_mod <- MVNIGmod(gaad_res[1:50], gaad_cov[1:50], initial_mvnig_theta)

predict(mvnig_mod, gaad_cov[1:50])
```

theta_mvnig *Toy Response Initial Parameter (MVNIG)*

Description

Part of toy dataset for examples.

Usage

```
theta_mvnig
```

Format

List of parameters for input to MVNIGmod function

Examples

```
theta_mvvg
```

theta_mvvg *Toy Response Initial Parameter (MVVG)*

Description

Part of toy dataset for examples.

Usage

```
theta_mvvg
```

Format

List of parameters for input to MVVGmod function

Examples

```
theta_mvvg
```

X *Toy Covariate Matrices*

Description

Part of toy dataset for examples.

Usage

```
X
```

Format

List of covariate matrices for individual subjects

Examples

```
X
```

Y *Toy Response Matrices*

Description

Part of toy dataset for examples.

Usage

```
Y
```

Format

List of response matrices for individual subjects

Examples

```
Y
```

Index

∗ **datasets**
gaad_cov, 2
gaad_res, 3
gaad_theta_mvvg, 3
theta_mvnig, 9
theta_mvvg, 9
X, 10
Y, 10

gaad_cov, 2
gaad_res, 3
gaad_theta_mvvg, 3

MVNIGmod, 4
MVVGmod, 6

predict, 8

theta_mvnig, 9
theta_mvvg, 9

X, 10

Y, 10