Package 'blin'

September 22, 2018

Title Bipartite Longitudinal Influence Network (BLIN) Estimation

Version 0.0.1

Description Estimate influence networks from longitudinal bipartite relational data, where the longitudinal relations are continuous. The outputs are estimates of weighted influence networks among each actor type in the data set. The generative model is the Bipartite Longitudinal Influence Network (BLIN) model, a linear autoregressive model for these type of data. The supporting paper is ``Inferring Influence Networks from Longitudinal Bipartite Relational Data", which is in preparation by the same authors. The model may be estimated using maximum likelihood methods and Bayesian methods. For more detail on methods, see Marrs et. al. <arXiv:1809.03439>.

Depends R (>= 3.3.0)

Imports glmnet, stats, Matrix, MASS, abind, graphics, mvtnorm

VignetteBuilder knitr

Suggests knitr, knitcitations

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

NeedsCompilation no

Author Frank W. Marrs [aut, cre],
Benjamin W. Campbell [aut],
Bailey K. Fosdick [aut],
Skyler J. Cranmer [aut],
Tobias B{"o}hmelt [aut]

Maintainer Frank W. Marrs <frank.marrs@colostate.edu>

Repository CRAN

Date/Publication 2018-09-21 22:30:10 UTC

R topics documented:

`blin_mle` *Estimate the BLIN model using maximum likelihood estimator*

Description

This function estimates the bipartite logitudinal influence network (BLIN) model $Y_t = A^T \sum_{k=1}^{lag} Y_{t-k} + \sum_{k=1}^{lag} Y_{t-k} B + X_t \beta + \tau E_t$ using maximum likelihood estimator.

Usage

```
blin_mle(Y, X = NULL, type = "full", lag = 1, rankA = NULL,
  rankB = rankA, maxit = 1000, tol = 1e-08, init = "I",
  sigma_init = 1, verbose = FALSE, calcses = FALSE, randseed = NA)
```

Arguments

`Y`	Response 3-mode array.
`X`	Optional 4-mode array of covariates, defaults to no covariates.
`type`	Optional string specifying BLIN model type: full, reduced_rank, or sparse. Defaults to full.
`lag`	Optional numeric specifying autoregressive lag in model, defaults to 1.
`rankA`	Optional numeric rank of influence network matrix A for reduced rank model type, defaults to full rank.
`rankB`	Optional numeric rank of influence network matrix B, defaults to rank of A.
`maxit`	Optional numeric maximum number of iterations for full and reduced rank block coordinate descents, defaults to 1e3.
`tol`	Optional numeric convergence tolerance for full and reduced rank block coordinate descents, defaults to 1e-8.

`init`	Optional string specifying initialization type for full and reduced rank block coordinate descents, defaults to "I", identity for A and B. Also allows "random" for random initialization of A and B.
`sigma_init`	Optional numeric standard deviation for random initialization of A and B in full and reduced rank block coordinate descents, defaults to 1.
`verbose`	Optional logical specifying whether progress should be printed out (`TRUE`) or not (`FALSE`). Defaults to `FALSE`.
`calcses`	Optional logical specifying whether standard errors should be calculated (`TRUE`) or not (`FALSE`). Defaults to `FALSE`. Only standard errors for the full BLIN model are implemented.
`randseed`	Optional numeric specifying seed for random ininitialization of A and B in full and reduced rank block coordinate descents, defaults to `NA` (no seed set).

Details

This function estimates the continuous BLIN model,

$$Y_t = A^T Y_{t-1} + Y_{t-1} B + X_t \beta + \tau E_t$$

, where $\{Y_t\}_t$ is a set of $S \times L$ matrices representing the bipartite relation data at each observation t. The set $\{X_t\}_t$ is a set of $S \times L \times p$ arrays describing the influence of the coefficient vector $beta$. Finally, each matrix E_t is assumed to consist of iid standard normal random variables. The matrices A and B are square matrices respesenting the influence networks among S senders and L receivers, respectively.

This function estimates the BLIN model using maximum likelihood (and related) methods. The "full" model places no restrications on the influence networks A and B, and estimates these matrices (along with β) by block coordinate descent. In addition, if `calcses==TRUE`, the standard errors for each coefficient will be estimated. Note that the standard error procedure may require large amounts of memory to build the BLIN design matrix; a warning is produced if the estimated size of the desgn is greater than 0.5GB.

The "reduced rank" BLIN model assumes that the matrix A has decomposition $A = UV^T$, where each of U and V is an $S \times$ `rankA` matrix, and the matrix B has decomposition $B = WZ^T$, where each of W and Z is an $L \times$ `rankB` matrix. This model is also estimated using block coordinate descent.

Finally, the "sparse" BLIN model assumes that A and B matrices have many entries that are small or zero. The `cv.glmnet(.)` function from the `glmnet` package is used to estimate the entries in A, B, and $beta$. The object resuling from `cv.glmnet(.)` is returned in this case.

Notice that the diagonals of A and B are not identifiable. However, the sum of each diagonal entry in A and B, i.e. $a_{ii} + b_{jj}$, is identifiable. Thus, the diagonal sums are broken out as separate estimates under the name `diagAB`.

If `calcses = TRUE` and `type = full`, then standard errors will be returned. These standard errors are based on the assumption that each E_t consists of iid standard normal random variables. In this case, the full design matrix is built, which we call W here. Then, the variance-covariance matrix of the estimated coefficients is formed by $\hat{\tau}^2 (W^T W)^{-1}$, where $\hat{\tau}^2$ is the usual unbiased estimator of the error variance.

Value

fit	A blin object containing summary information.

See Also

generate_blin build_design

Examples

```
S <- 5
L <- 4
tmax <- 10
data <- generate_blin(S,L,tmax, lag=2, sparse=.8, seed=1)

fit <- blin_mle(data$Y, data$X, lag=2, calcses=TRUE)
summary(fit)
```

---

build_design *Build the BLIN design matrix*

Description

Build the BLIN design matrix

Usage

```
build_design(Y, X = NULL, lag = 1, showWarnings = TRUE)
```

Arguments

Y	Response 3-mode array.
X	Optional 4-mode array of covariates, defaults to no covariates.
lag	Optional numeric specifying autoregressive lag in model, defaults to 1.
showWarnings	Optional logical whether matrix memory size should be evaluated and warning provided (see details), defaults to TRUE.

Details

This function takes an $S \times L \times T$ array Y that is a representation of a longitudinal bipartite relational data set. Optional input is an $S \times L \times T \times p$ array X of covariates that influence the evolution of the data set in equation over time. The function returns an $(SL(T - lag)) \times (S^2 + L^2 + p)$ design matrix, of sparse class, upon which `Y[,,lag:T]` may be regressed. If `showWarnings = TRUE`, and if the estimated size of the design matrix is greater than 1GB, a warning is thrown.

Value

A sparse design matrix

See Also

generate_blin blin_mle

Examples

```
S <- 5
L <- 4
tmax <- 10
data <- generate_blin(S,L,tmax, lag=2, sparse=.8, seed=1)
dim(data$Y)

Xreg <- build_design(data$Y, data$X, lag=2)
dim(Xreg)
class(Xreg)
```

---

coef.blin *Coef S3 generic for class blin*

Description

Coef S3 generic for class blin

Usage

```
## S3 method for class 'blin'
coef(object, whichcoef = NULL, ...)
```

Arguments

object	blin object
whichcoef	optional string (or NULL) indicating which coefficient to retrun, i.e. A, B, beta, or diagAB. If NULL, returns list of all coefficients.
...	ignored

---

forum *Online forum dataset*

Description

A data set containing online forum posts from students at the University of California at Irvine, from 2004 (see Opsahl 2013).

Format

A data set with a single array

forum 20 x 20 x 24 numeric matrix of weights. NA at (i, j, t) indicates that user i did not post to forum j in week t.

Details

This data set contains online forum posts from students at the University of California at Irvine, from 2004 (see Opsahl 2013). The 20 most active users and the 20 forums to which these users posted the most are examined. The weights of the network are the number of characters posted to a given forum by a given user for each week. The 3-mode array `forum` contains the weights indexed by user, forum, and week, respectively. Data obtained June 8, 2018. See the link http://opsahl.co.uk/tnet/datasets/OF_longitudinal_weightedchar.txt for raw data.

Source

http://opsahl.co.uk/tnet/datasets/OF_longitudinal_weightedchar.txt

References

Opsahl, T. (2013). "Triadic closure in two-mode networks: Redefining the global and local clustering coefficients." Social Networks, 35(2), 159-167. <doi:10.1016/j.socnet.2011.07.001>

Examples

```
data("forum")
```

`generate_blin` *Generate data from the continuous BLIN model*

Description

This function generates data from the bipartite logitudinal influence network (BLIN) model $Y_t = A^T \sum_{k=1}^{lag} Y_{t-k} + \sum_{k=1}^{lag} Y_{t-k} B + X_t \beta + \tau E_t$.

Usage

```
generate_blin(S, L, tmax, lag = 1, tau = 1, sigmaY = 1, muAB = 0,
  sigmaAB = 1, rankA = S, rankB = L, use_cov = TRUE, seed = NA,
  sparse = NA)
```

Arguments

S	Dimension of A.
L	Dimension of B.
tmax	Number of observations of relational data.
lag	Autoregressive lag in model, defaults to 1.
tau	Optional error standard deviatiom, defaults to 1.
sigmaY	Optional standard deviation of entries in Y_t, defaults to 1.
muAB	Optional mean of entries in decomposition of matrices $A = UV^T$ and $B = WZ^T$, defaults to 0.
sigmaAB	Optional standard deviation of entries in decomposition matrices of $A = UV^T$ and $B = WZ^T$, defaults to 1.
rankA	Rank of influence network matrix A, defaults to full rank.
rankB	Optional rank of influence network matrix B, defaults to full rank.
use_cov	Optional logical used to indicate whether to include $X_t\beta$ in the model (TRUE) or not (FALSE), defaults to TRUE.
seed	Optional numeric to set seed before generating, defaults to NA (no seed set).
sparse	Optional degree of sparsity in A and B, i.e. sparsity=.9 means 10% of the entries in A and B are set to zero at random. Defaults to NA (no entries set to zero).

Details

This function generates a continuous bipartite longitudinal relational data set from the BLIN model, $Y_t = A^T \sum_{k=1}^{lag} Y_{t-k} + \sum_{k=1}^{lag} Y_{t-k}B + X_t\beta + \tau E_t$, where $\{Y_t\}_t$ is a set of $S \times L$ matrices representing the bipartite relational data at each observation t. The set $\{X_t\}_t$ is a set of $S \times L \times p$ arrays describing the influence of the coefficient vector $beta$. Finally, each matrix E_t consists of iid standard normal random variables.

The matrices A and B are square matrices respesenting the influence networks among S senders and L receivers, respectively. The matrix A has decomposition $A = UV^T$, where each of U and V is an $S \times rankA$ matrix of iid standard normal random variables with mean muAB and standard deviation sigmaAB. Similarly, the matrix B has decomposition $B = WZ^T$, where each of W and Z is an $L \times rankB$ matrix of iid standard normal random variables with standard deviation sigmaAB and mean muAB for W and mean -muAB for Z. Lastly, the covariate array X_t has 3 covariates: the first is an intercept, the second consists of iid Bernoulli random variables, and the third consists of iid standard normal random variables. All coefficients are $\beta_i = 0$ for $i = 1, 2, 3$.

Value

fit	An blin object containing summary information.

See Also

blin_mle

Examples

```
S <- 5
L <- 4
tmax <- 10
data <- generate_blin(S,L,tmax, lag=2, sparse=.8)
names(data)
dim(data$X)
data$A
```

`model.matrix.blin` *model.matrix S3 generic for class blin*

Description

model.matrix S3 generic for class blin

Usage

```
## S3 method for class 'blin'
model.matrix(object, ...)
```

Arguments

`object`	blin object
`...`	ignored

`plot.blin` *Plot S3 generic for class blin*

Description

Plot S3 generic for class blin

Usage

```
## S3 method for class 'blin'
plot(x, ...)
```

Arguments

`x`	blin object
`...`	ignored

print.blin *Print S3 generic for class blin*

Description

Print S3 generic for class blin

Usage

```
## S3 method for class 'blin'
print(x, hn = 10, ...)
```

Arguments

x	blin object
hn	optional numeric length of each coefficient printed
...	ignored

print.summary.blin *Print S3 generic for class summary.blin*

Description

Print S3 generic for class summary.blin

Usage

```
## S3 method for class 'summary.blin'
print(x, hn = 10, ...)
```

Arguments

x	summary.blin object
hn	optional numeric length of each coefficient printed
...	ignored

summary.blin *Summary S3 generic for class blin*

Description

Summary S3 generic for class blin

Usage

```
## S3 method for class 'blin'
summary(object, whichcoef = NULL, ...)
```

Arguments

object	blin object
whichcoef	optional string (or NULL) indicating which coefficient to retrun, i.e. A, B, beta, or diagAB. If NULL, returns list of all coefficients.
...	ignored

vcov.blin *vcov S3 generic for class blin*

Description

vcov S3 generic for class blin

Usage

```
## S3 method for class 'blin'
vcov(object, ...)
```

Arguments

object	blin object
...	ignored

Index