Package ‘PAGE’

July 21, 2025

Type Package

Title Predictor-Assisted Graphical Models under Error-in-Variables

Version 0.1.0

Description We consider the network structure detection for variables Y with auxiliary variables X accommodated, which are possibly subject to measurement error. The following three functions are designed to address various structures by different methods : one is NP_Graph() that is used for handling the nonlinear relationship between the responses and the covariates, another is Joint_Gaussian() that is used for correction in linear regression models via the Gaussian maximum likelihood, and the other Cond_Gaussian() is for linear regression models via conditional likelihood function.

License GPL-3

Encoding UTF-8

Imports glasso, lars, network, GGally, caret, randomForest, metrica, MASS, stats

Suggests sna

RoxygenNote 7.3.2

Author Wan-Yi Chang [aut, cre],
Li-Pang Chen [aut]

Maintainer Wan-Yi Chang <jessica306a@gmail.com>

NeedsCompilation yes

Repository CRAN

Date/Publication 2025-07-21 08:42:18 UTC

Contents

Cond_Gaussian — *Estimation of network structure and variable selection in the linear model via the conditional likelihood function.*

Description

This function focuses on multivariate linear regression models $Y = XB + \epsilon$ subject to measurement error in responses and covariates, where with B is a matrix of parameters and ϵ is a noise term with zero expectation. We aim to detect the network structure of responses and select informative covaraites. The estimation method is motivated by the conditional likelihood function and uses the conditional expectation to correct for measurement error.

Usage

```
Cond_Gaussian(
  W,
  Z,
  sigma_eta,
  sigma_delta,
  alpha_1,
  alpha_2,
  alpha_1_list = NULL,
  alpha_2_list = NULL,
  max_iter = 30,
  tol = 1e-06,
  label_name = TRUE
)
```

Arguments

W	A $n \times m$ response matrix, the variables can be error-prone or precisely measured.
Z	A $n \times p$ matrix of continuous covariates. The variables can be error-prone or precisely measured.
sigma_eta	A $p \times p$ covariance matrix of the noise term η in the classical measurement error model $Z = X + \eta$, where X is the unobserved version of Z.
sigma_delta	A $m \times m$ covariance matrix of the noise term δ in the classical measurement error model $W = Y + \delta$, where Y is the unobserved version of W.
alpha_1	A tuning parameter associated with parameter B.
alpha_2	A tuning parameter associated with parameter, denoted as Γ, that reflects the network in Y.
alpha_1_list	A list of tuning parameters for the model averaging estimator of B. The default value is NULL.
alpha_2_list	A list of tuning parameters for the model averaging estimator of Γ. The default value is NULL.

`max_iter`	A maximum number for iterations for updated values of B and Γ. The default value is 30.
`tol`	A prespecified tolerance ζ for iterations for updated values of B and Γ. The default value is 10^{-6}.
`label_name`	The name of the response variables. The default value is TRUE, which reflects the labels from the input data. Else, users can input the required labels manually.

Value

`B`	An estimator of B.
`gamma`	An estimator of the network in Y.
`graph`	A visualization of the estimated network structure by gamma.
`Beta_BICs`	A vector of Bayesian Information Criterion (BIC) weights for the model averaging estimator of B under candidate models alpha_1_list.
`Gamma_BICs`	A vector of Bayesian Information Criterion (BIC) weights for the model averaging estimator of Γ under candidate models alpha_2_list.

Author(s)

Wan-Yi Chang and Li-Pang Chen
Maintainer: Wan-Yi Chang <`jessica306a@gmail.com`>

Examples

```
n <- 100
Z <- matrix(rnorm(n * 5), n, 5)
W <- matrix(rnorm(n * 5), n, 5)
sigma_eta <- diag(0.15, ncol(Z))
sigma_delta <- diag(0.3, ncol(W))

Cond_Gaussian(W, Z, sigma_eta, sigma_delta,
                        alpha_1 = 0.1, alpha_2 = 0.1,
                        alpha_1_list = NULL,
                        alpha_2_list = NULL,
                        max_iter = 1, tol = 1e-6, label_name = TRUE)
```

`Joint_Gaussian` *Estimation of network structure and variable selection in the linear model via the Gaussian maximum likelihood.*

Description

This function focuses on multivariate linear regression models $Y = XB + \epsilon$ subject to measurement error in the responses and covariates, where with B is a matrix of parameters and ϵ is a noise term with zero expectation. We aim to detect the network structure of responses and select informative covaraites. The estimation method is motivated by the Gaussian maximum likelihood function and uses the conditional expectation to correct for measurement error.

Usage

```
Joint_Gaussian(
  W,
  Z,
  sigma_eta,
  sigma_delta,
  alpha_1,
  alpha_2,
  alpha_1_list = NULL,
  alpha_2_list = NULL,
  label_name = TRUE
)
```

Arguments

W	A $n \times m$ response matrix, the variables can be error-prone or precisely measured.
Z	A $n \times p$ matrix of continuous covariates. The variables can be error-prone or precisely measured.
sigma_eta	A $p \times p$ covariance matrix of the noise term η in the classical measurement error model $Z = X + \eta$, where X is the unobserved version of Z.
sigma_delta	A $m \times m$ covariance matrix of the noise term δ in the classical measurement error model $W = Y + \delta$, where Y is the unobserved version of W.
alpha_1	A tuning parameter associated with the parameter B.
alpha_2	A tuning parameter associated with precision matrix C, which is the inverse of the covariance matrix of ϵ.
alpha_1_list	A list of tuning parameters for the model averaging estimator of B. The default value is NULL.
alpha_2_list	A list of tuning parameters for the model averaging estimator of C. The default value is NULL.
label_name	The name of the response variable. The default value is TRUE, which reflects the labels from the input data. Else, users can input the required labels manually.

Value

B	An estimator of B.
C	An estimator of C.
graph	A visualization of the estimated network structure by C.

`Beta_BICs` A vector of Bayesian Information Criterion (BIC) weights for the model averaging estimator of B under candidate models alpha_1_list.

`Gamma_BICs` A vector of Bayesian Information Criterion (BIC) weights for the model averaging estimator of C under candidate models alpha_2_list.

Author(s)

Wan-Yi Chang and Li-Pang Chen
Maintainer: Wan-Yi Chang <jessica306a@gmail.com>

Examples

```
n <- 100
Z <- matrix(rnorm(n * 10), n, 10)
W <- matrix(rnorm(n * 5), n, 5)
sigma_eta <- diag(0.15, ncol(Z))
sigma_delta <- diag(0.3, ncol(W))

Joint_Gaussian(W, Z, sigma_eta, sigma_delta,
                         alpha_1 = 0.1, alpha_2 = 0.1,
                         alpha_1_list = c(0.1, 0.3),
                         alpha_2_list = c(0.1, 0.3),
                         label_name = TRUE)
```

`NP_Graph` *Estimation of network structure and variable selection in the nonlinear model with measurement errors in responses and covariates.*

Description

This function characterizes Y and X by nonlinear models and is designed for detecting network structure and variable selection with measurement error in responses and covariates. Here the components of Y can be continuous, binary, or count. The estimation strategy in this function includes the regression calibration for correcting error-prone responses and covariates, the random forest method for marginally characterizing the response and covariates, and the distance correlation and graphical lasso for detecting the network structure among the responses.

Usage

```
NP_Graph(
  W,
  Z,
  sigma_eta,
  rho,
  sigma_delta = 0.5,
  r = 0.8,
  lambda = 1,
```

```
  pi = 0.8,
  label_name,
  var_thred = 5
)
```

Arguments

W	A n × m response matrix. The variables can be error-prone or precisely measured, and can include continuous, binary, or count random variables.
Z	A n × p matrix of continuous covariates. The variables can be error-prone or precisely measured.
sigma_eta	A p × p covariance matrix of the noise term η in the classical measurement error model Z = X + η, where X is the unobserved version of Z.
rho	A tuning parameter for the graphical lasso.
sigma_delta	The common value in the diagonal covariance matrix of the noise term δ in the classical measurement error model for continuous components in W. The default value is 0.5.
r	A probability r for misclassification when components in W are binary. The default value is 0.8.
lambda	A parameter λ in the Poisson distribution that provides the increasing measurement error effects when components in W are count. The default value is 1.
pi	A parameter π in [0,1] for the Binomial distribution that characterizes the decreasing measurement error effects when components in W are count. The default value is 0.8.
label_name	The name of the response variable. The default value is TRUE, which reflects the labels from the input data. Else, users can input the required labels manually.
var_thred	A positive value used to retain important covariates. That is, covariates will be selected when refitting the model if their importance scores are greater than var_thred. The default value is 5.

Value

W_hat	The n × m matrix of corrected responses determined by regression calibration.
Z_hat	The n × p matrix of corrected covariates determined by regression calibration..
PSE	The Frobenius norm of the residual corresponding to W_hat.
importance_score	A matrix containing importance scores for the covariates.
precision_matrix	An estimated matrix reflecting the network structure of the responses.
graph	An visualization of the estimated network structure by `precision_matrix`.

Author(s)

Wan-Yi Chang and Li-Pang Chen
Maintainer: Wan-Yi Chang <jessica306a@gmail.com>

Examples

```
n <- 100
Z <- matrix(rnorm(n * 10), n, 10)
W <- matrix(rnorm(n * 5), n, 5)
sigma_eta <- diag(0.15, ncol(Z))

NP_Graph(W, Z, sigma_eta, rho = 0.2,
                   sigma_delta = 0.5, r = 0.8,
                   lambda = 1, pi = 0.8,
                   label_name = TRUE, var_thred = 3)
```

`PAGE_package` *Predictor-Assisted Graphical Models under Error-in-Variables*

Description

This package has three functions that characterize the multivariate responses and covariates under linear or nonlinear structures. All functions are valid to handle measurement error, detection of network in responses, and selection of informative covariates.

Details

Given the responses and covariates that can be error-prone or precisely measured, NP_graph is a function used to detect the network structure of responses and select important covariates under nonlinear structures. Under linear structures, this package provides two different estimation methods: the function Joint_Gaussian implements the error-corrected Gaussian maximum likelihood method, and the function Cond_Gaussian extends the neighborhood selection strategy to construct the corrected conditional likelihood function. All functions are able to estimate the network strucure of responses and perform variable selection to identify important covariates with measurement error taken into account.

Value

PAGE_package

Index