Package ‘RidgeVar’

September 11, 2020

Type Package

Title Estimation of error variance via ridge regression

Version 1.0.1

Author Xu Liu [aut,cre],
Shurong Zheng [aut],
Xingdong Feng [aut],
Xiao Zhang [ctb]

Maintainer Xu Liu <liu.xu@sufe.edu.cn>

Description Provide several methods to estimate the error variance for high-dimensional linear regression models, which includes the ridge regression based method of Liu et al. (2020), the refitted cross validation of Fan et al. (2012), the maximum likelihood based method of Dicker and Erdogdu (2016), and the moments based method of Dicker (2014).

License GPL (>= 2)

Depends R (>= 3.2.0)

Imports glmnet

LazyData true

NeedsCompilation yes

Repository CRAN

URL https://github.com/xliusufe/RidgeVar

Encoding UTF-8

Archs i386, x64

R topics documented:

`RidgeVar-package` *Estimation of error variance via ridge regression*

Description

Provide several methods to estimate the error variance for high-dimensional linear regression models, which includes the ridge regression based method of Liu et al. (2020), the refitted cross validation of Fan et al. (2012), the maximum likelihood based method of Dicker and Erdogdu (2016), and the moments based method of Dicker (2014).

Details

Package:	RidgeVar
Type:	Package
Version:	1.0.1
Date:	2019-03-12
License:	GPL (>= 2)

References

Dicker, L. H. (2014). Variance estimation in high-dimensional linear models. Biometrika 101, 269-284.

Dicker, L. H. and Erdogdu, M. A. (2016). Maximum likelihood for variance estimation in high-dimensional linear models. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS 2016), 159-167. JMLR Workshop & Conference Proceedings.

Fan, J., Guo, S. and Hao, N. (2012). Variance estimation using refitted cross-validation in ultrahigh-dimensional regression. Journal of Royal Statistical Society Series B 74, 37-65.

Liu, X., Zheng, S. and Feng, X. (2020). Estimation of error variance via ridge regression. Biometrika, 107, 481-488.

`COV_RR` *Ridge Regression Based Covariance Estimation in High-Dimensional Multivariate Linear Models*

Description

Use the ridge regression based method of Liu et al. (2020) to estimate the residual covariance in multivariate linear model. The proposal is valid under both low- and high-dimensional models, and performs well not only for the non-sparse cases but also for the sparse ones.

Usage

```
COV_RR(y,x,eta=NULL,alpha=0.1)
```

Arguments

y	A $n \times q$ response matrix.
x	A $n \times p$ numeric design matrix for the model.
eta	The tunning parameter using in ridge regression. Default is NULL, and `eta` will be selected by the method proposed in Liu et al. (2020) in this case.
alpha	A constant to justify the tunning parameter when eta=NULL. Default is 0.1.

Value

sigma2	The estimation of the residual covariance.
trA1	The value of $n - \text{tr}(A_{1n})$, see the details in Liu, X., Zheng, S. and Feng, X. (2020).

References

Liu, X., Zheng, S. and Feng, X. (2020). Estimation of error variance via ridge regression. Biometrika, 107, 481-488.

Examples

```
n   <- 80
p   <- 100
q   <- 5
eta <- 1e-6
rho <- 0.8
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
sig = chol( outer(1:q,1:q,function(i,j) rho^abs(i-j)) )
y <- matrix(rnorm(n*q),n,q) %*% sig
x <- matrix(rnorm(n*p),n,p)
y[,1] <- y[,1] + x %*% beta
fit <- COV_RR(y,x,eta)
```

VAR_MLE — *Likelihood Based Variance Estimation in High-Dimensional Linear Models*

Description

Use the maximum likelihood based method of Dicker and Erdogdu (2016) to estimate the residual variance in high-dimensional linear model.

Usage

```
VAR_MLE(y,x,max.iter=50,tol=1e-4)
```

Arguments

y	A length n vector of response.
x	A $n \times p$ numeric design matrix for the model.
max.iter	Maximum number of iterations. Default is 50.
tol	Convergence threshhold. Default is 1e-4.

Value

sigma2	The estimation of the residual variance.

References

Dicker, L. H. and Erdogdu, M. A. (2016). Maximum likelihood for variance estimation in high-dimensional linear models. In Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS 2016), 159-167. JMLR Workshop & Conference Proceedings.

Examples

```
n   <- 80
p   <- 100
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
x <- matrix(rnorm(n*p),n,p)
y <- rnorm(n)
y <- y + x %*% beta
fit <- VAR_MLE(y,x)
```

VAR_MM *Moments Based Variance Estimation in High-Dimensional Linear Models*

Description

Use the moments based method of Dicker (2014) to estimate the residual variance in high-dimensional linear model.

Usage

```
VAR_MM(y,x,identity=F,Sigma=NULL)
```

Arguments

y	A length n vector of response.
x	A $n \times p$ numeric design matrix for the model.
identity	Logical indicating that the covariance of X is identity if identity=TRUE, and not if identity=FALSE. Default is FALSE. It is invalid if Sigma is not NULL.
Sigma	A $p \times p$ matrix, which is the covariance of X. If Sigma=NULL, the sample covarince is given. Default is NULL.

Value

sigma2	The estimator of the residual variance.

References

Dicker, L. H. (2014). Variance estimation in high-dimensional linear models. Biometrika **101**, 269-284.

Examples

```
n   <- 80
p   <- 100
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
eps <- rnorm(n)
x <- matrix(rnorm(n*p),n,p)
y <- x%*%beta+eps
fit <- VAR_MM(y,x)
```

`VAR_RCV` *Refitted Cross-Validation Based Variance Estimation in High-Dimensional Linear Models*

Description

Use the refitted cross-validation based method of Fan et al. (2012) to estimate the residual variance in high-dimensional linear model.

Usage

```
VAR_RCV(y,x)
```

Arguments

`y`	A length n vector of response.
`x`	A $n \times p$ numeric design matrix for the model.

Value

`sigma2`	The estimation of the residual variance.

References

Fan, J., Guo, S. and Hao, N. (2012). Variance estimation using refitted cross-validation in ultrahigh-dimensional regression. Journal of Royal Statistical Society, Series B **74**, 37-65.

Examples

```
n   <- 80
p   <- 100
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
eps <- rnorm(n)
x <- matrix(rnorm(n*p),n,p)
y <- x%*%beta+eps
fit <- VAR_RCV(y,x)
```

VAR_RR *Ridge Regression Based Variance Estimation in High-Dimensional Linear Models*

Description

Use the ridge regression based method of Liu et al. (2020) to estimate the residual variance in (multivariate) linear model. The proposal is valid under both low- and high-dimensional models, and performs well not only for the non-sparse cases but also for the sparse ones.

Usage

```
VAR_RR(y,x,eta=NULL,alpha=0.1)
```

Arguments

y	A $n \times q$ response matrix. y is an n-vector if $q = 1$.
x	A $n \times p$ numeric design matrix for the model.
eta	The tunning parameter using in ridge regression. Default is NULL, and `eta` will be selected by the method proposed in Liu et al. (2020) in this case.
alpha	A constant to justify the tunning parameter when eta=NULL. Default is 0.1.

Value

sigma2	The estimation of the residual variance, which is a q vector.
trA1	The value of $n - \text{tr}(A_{1n})$, see the details in Liu, X., Zheng, S. and Feng, X. (2020).

References

Liu, X., Zheng, S. and Feng, X. (2020). Estimation of error variance via ridge regression. Biometrika, 107, 481-488.

Examples

```
# example 1
n   <- 80
p   <- 100
eta <- 1e-6
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
x <- matrix(rnorm(n*p),n,p)
y <- rnorm(n)
y <- y + x%*%beta
fit <- VAR_RR(y,x,eta)

# example 2
n   <- 80
p   <- 100
q   <- 5
eta <- 1e-6
rho <- 0.8
```

```
beta <- c(sqrt(0.1/p)*rep(1,p/2),rep(0,p/2))
sig = chol( outer(1:q,1:q,function(i,j) rho^abs(i-j)) )
y <- matrix(rnorm(n*q),n,q) %*% sig
x <- matrix(rnorm(n*p),n,p)
y[,1] <- y[,1] + x %*% beta
fit <- VAR_RR(y,x,eta)
```

Index