

Package ‘GWRLASSO’

August 28, 2023

Type Package

Title A Hybrid Model for Spatial Prediction Through Local Regression

Version 0.1.0

Depends R(>= 2.10)

Suggests knitr, rmarkdown, testthat

Description It implements a hybrid spatial model for improved spatial prediction by combining the variable selection capability of LASSO (Least Absolute Shrinkage and Selection Operator) with the Geographically Weighted Regression (GWR) model that captures the spatially varying relationship efficiently. For method details see, Wheeler, D.C.(2009).<[DOI:10.1068/a40256](https://doi.org/10.1068/a40256)>. The developed hybrid model efficiently selects the relevant variables by using LASSO as the first step; these selected variables are then incorporated into the GWR framework, allowing the estimation of spatially varying regression coefficients at unknown locations and finally predicting the values of the response variable at unknown test locations while taking into account the spatial heterogeneity of the data. Integrating the LASSO and GWR models enhances prediction accuracy by considering spatial heterogeneity and capturing the local relationships between the predictors and the response variable. The developed hybrid spatial model can be useful for spatial modeling, especially in scenarios involving complex spatial patterns and large datasets with multiple predictor variables.

License GPL (>= 2.0)

Encoding UTF-8

Imports stats, qpdf, numbers, glmnet, Matrix

RoxygenNote 7.2.3

VignetteBuilder knitr

NeedsCompilation no

Author Nobin Chandra Paul [aut, cre, cph],
Anil Rai [aut],
Ankur Biswas [aut],
Tauqueer Ahmad [aut],

Bhaskar B. Gaikwad [aut],
 Dhananjay D. Nangare [aut],
 K. Sammi Reddy [aut]

Maintainer Nobin Chandra Paul <nobin.paul@icar.gov.in>

Repository CRAN

Date/Publication 2023-08-28 11:10:05 UTC

R topics documented:

GWRLASSO_exponential	2
GWRLASSO_gaussian	3
Index	6

GWRLASSO_exponential *GWRLASSO: a hybrid model that uses the LASSO model for important variable selection and GWR model with exponential kernel for prediction at an unknown location based on the selected variables.*

Description

GWRLASSO: a hybrid model that uses the LASSO model for important variable selection and GWR model with exponential kernel for prediction at an unknown location based on the selected variables.

Usage

```
GWRLASSO_exponential(data_sp, bw, split_value, exponential_kernel, nfolds)
```

Arguments

data_sp	A dataframe containing the response variable and the predictor variables, as well as the coordinates of the locations. In the dataframe, first column is the response variable (y), last two columns are coordinates i.e., Latitude and Longitudes and in between them is the set of predictor variables(X's).
bw	A numeric value specifying the bandwidth parameter for the GWR model. It can be noted that, optimum bandwidth value can vary depending on the specific dataset and bandwidth parameter depends on the spatial pattern of the data
split_value	Splitting value for dividing the dataset into training and testing sets, e.g. 0.7 or 0.8
exponential_kernel	Spatial weight function of the GWR model, e.g. exponential_kernel
nfolds	Specifies the number of folds to use in the cross-validation process for selecting the optimal value of the regularization parameter in LASSO

Value

A list with the following components: - 'Important_vars ': Selected important variables based on LASSO model - 'Optimum_lamda': Optimum lamda value obtained by CV approach - 'GWR_y_pred_test': The GWR predictions at testing locations - 'R_square': R-square value - 'rrmse': relative root means square error value - 'mse': mean squared error - 'mae': mean absolute error

References

1. Brunson, C., Fotheringham, A.S. and Charlton, M.E. (1996). Geographically weighted regression: a method for exploring spatial non-stationarity. *Geogr Anal.*28(4),281-298.<DOI:10.1111/j.1538-4632.1996.tb00936.x>. 2. Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*,33(1),1-22.<DOI:10.18637/jss.v033.i01>. 3. Wheeler, D. C. (2009). Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and planning A*, 41(3), 722-742.<DOI:10.1068/a40256>.

Examples

```
n<- 100
p<- 7
m<-sqrt(n)
id<-seq(1:n)
x<-matrix(runif(n*p), ncol=p)
e<-rnorm(n, mean=0, sd=1)
xy_grid<-expand.grid(c(1:m),c(1:m))
Latitude<-xy_grid[,1]
Longitude<-xy_grid[,2]
B0<-(Latitude+Longitude)/6
B1<-(Latitude/3)
B2<-(Longitude/3)
B3<-(2*Longitude)
B4<-2*(Latitude+Longitude)/6
B5<-(4*Longitude/3)
B6<-2*(Latitude+Longitude)/18
B7<-(4*Longitude/18)
y<-B0+(B1*x[,1])+(B2*x[,2])+(B3*x[,3])+(B4*x[,4])+(B5*x[,5])+(B6*x[,6])+(B7*x[,7])+e
data_sp<-data.frame(y,x,Latitude,Longitude)
GWRLASSO_exp<-GWRLASSO_exponential(data_sp,0.8,0.7,exponential_kernel,10)
```

GWRLASSO_gaussian

GWRLASSO: a hybrid model that uses the LASSO model for important variable selection and GWR model with gaussian kernel for prediction at an unknown location based on the selected variables.

Description

GWRLASSO: a hybrid model that uses the LASSO model for important variable selection and GWR model with gaussian kernel for prediction at an unknown location based on the selected variables.

Usage

```
GWRLASSO_gaussian(data_sp, bw, split_value, gaussian_kernel, nfolds)
```

Arguments

data_sp	A dataframe containing the response variable and the predictor variables, as well as the coordinates of the locations. In the dataframe, first column is the response variable (y), last two columns are coordinates i.e., Latitude and Longitudes and in between them is the set of predictor variables(X's).
bw	A numeric value specifying the bandwidth parameter for the GWR model. It can be noted that, optimum bandwidth value can vary depending on the specific dataset and bandwidth parameter depends on the spatial pattern of the data
split_value	Splitting value for dividing the dataset into training and testing sets, e.g. 0.7 or 0.8
gaussian_kernel	Spatial weight function of the GWR model, e.g. gaussian_kernel
nfolds	Specifies the number of folds to use in the cross-validation process for selecting the optimal value of the regularization parameter in LASSO

Value

A list with the following components: - 'Important_vars ': Selected important variables based on LASSO model - 'Optimum_lambda': Optimum lambda value obtained by CV approach - 'GWR_y_pred_test': The GWR predictions at testing locations - 'R_square': R-square value - 'rrmse': relative root means square error value - 'mse': mean squared error - 'mae': mean absolute error

References

1. Brunson, C., Fotheringham, A.S. and Charlton, M.E. (1996). Geographically weighted regression: a method for exploring spatial non-stationarity. *Geogr Anal.*28(4),281-298.<DOI:10.1111/j.1538-4632.1996.tb00936.x>. 2. Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*,33(1),1-22.<DOI:10.18637/jss.v033.i01>. 3. Wheeler, D. C. (2009). Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and planning A*, 41(3), 722-742.<DOI:10.1068/a40256>.

Examples

```
n<- 100
p<- 7
m<-sqrt(n)
id<-seq(1:n)
x<-matrix(runif(n*p), ncol=p)
e<-rnorm(n, mean=0, sd=1)
xy_grid<-expand.grid(c(1:m),c(1:m))
Latitude<-xy_grid[,1]
Longitude<-xy_grid[,2]
B0<-(Latitude+Longitude)/6
B1<-(Latitude/3)
```

```
B2<-(Longitude/3)
B3<-(2*Longitude)
B4<-2*(Latitude+Longitude)/6
B5<-(4*Longitude/3)
B6<-2*(Latitude+Longitude)/18
B7<-(4*Longitude/18)
y<-B0+(B1*x[,1])+(B2*x[,2])+(B3*x[,3])+(B4*x[,4])+(B5*x[,5])+(B6*x[,6])+(B7*x[,7])+e
data_sp<-data.frame(y,x,Latitude,Longitude)
GWRLASSO_gau<-GWRLASSO_gaussian(data_sp,0.8,0.7,gaussian_kernel,10)
```

Index

GWRLASSO_exponential, [2](#)
GWRLASSO_gaussian, [3](#)