Package 'SpatialDimExp'

December 11, 2020

Type Package

Title Dimension Expansion Methods for Nonstationary Spatial Random Process

Version 0.1.0

Author Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu

Maintainer Shanshan Qin <ssqin267@yorku.ca>

Description
Implement dimension expansion methods for modelling nonstationary spatial random processes. The dimension expansion method initially proposed by Bornn et al.(2012) aims to find a dimensionally sparse projection where the originally nonstationary field exhibits stationarity.

License GPL (>= 2)

Encoding UTF-8

Imports Matrix (>= 1.2-18),
fields (>= 11.6),
matrixcalc (>= 1.0-3),
pracma (>= 2.2.9),
data.table (>= 1.13.2)

LazyData true

RoxygenNote 7.1.1

Suggests knitr,
rmarkdown

VignetteBuilder knitr

R topics documented:

`DimExpansion` *Main fitting procedure.*

Description

The fitting procedure is to minimize one of the objective functions, $f_{OLS}, f_{WLS}, f_{GLS}$. Its purpose is to estimate the semivariogram function and the coordinates of the learned dimensions. The method applied in each iteration is BFGS (Broyden,1979).

Usage

```
DimExpansion(X, Y, lambda, Maxd, EmpSemivariogram, para.initial,
  Method = "OLS", model = "Exponential")
```

Arguments

`X`	The records of location coordinates, a $n \times d$ matrix, where n is the number of locations, and d is the dimension of the record location. Usually, $d = 2$ (latitude and longitude coordinates).
`Y`	Realizations of the Gaussian process $Y(X)$.
`lambda`	A prespecified tuning parameter, $\lambda > 0$.
`Maxd`	A prespecified maximum coordinate dimension of each location.
`EmpSemivariogram`	The empirical semivariogram that is computed by `EmpSemivariogramFit`.
`para.initial`	Initial parameters of (ϕ, Z) in the objective functions, $f_{OLS}, f_{WLS}, f_{GLS}$.
`Method`	Three optional methods, Method = c("OLS","WLS","GLS"), which corresponds to the Least-squares, weighted least-squares, and generalized least-squares fitting procedures (default:OLS).
`model`	Semivariogram model, it has three options, "Exponential", "Power","Gaussian" (default: "Exponential").

Value

Return the estimates of ϕ and locations coordinates Z of the learned dimensions.

References

- Bornn, L.,Shaddick, G., and Zidek, J.V. (2012). Modeling non-stationary processes through dimension expansion. Journal of American Statistical Association, 107:281–289.
- Broyden, C.G.(1979). The convergence of a class of double-rank minimization algorithms. Journal of the Institute of Mathematics and Its Application,6:76-90.
- Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu. (2020) Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

Examples

```
n = 10  # number of locations.
Maxd = 3
SimRep = 1000 # number of observations for each location
lambda = 0.01
trueTheta = 10
set.seed(123)
Data = GenerateData(n=n, SimRep = SimRep,trueTheta = trueTheta)
Y = Data$YObs
X = Data$LocIndex[,1:2]
EmpSemivariogram = EmpSemivariogramFit(Y)
para.initial = c(1,1,Data$LocIndex[,-c(1:2)])
DimExpansion(X,Y,lambda,Maxd,EmpSemivariogram, para.initial, Method="OLS",model = "Exponential")
DimExpansion(X,Y,lambda,Maxd,EmpSemivariogram, para.initial, Method="GLS",model = "Exponential")
DimExpansion(X,Y,lambda,Maxd,EmpSemivariogram, para.initial, Method="WLS",model = "Exponential")
```

`EmpSemivariogramFit` *Estimate the empirical semivariogram function.*

Description

This is to estimate the empirical semivariogram function.

Usage

```
EmpSemivariogramFit(Y)
```

Arguments

`Y`	Realizations of the spatial random process $Y(X)$;

Value

A vector formed by concatenating the columns of upper triangular of the empirical semivariogram.

References

Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu.(2020) Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

`Gamma_phiFun` *The Isotropic Semivariogram Model.*

Description

Define a semivariogram function (no nugget effect). Herein, we provide three options, "Exponential", "Power","Gaussian".

Usage

```
Gamma_phiFun(phi, x, model = "Exponential")
```

Arguments

phi	Two dimensional vector-valued parameters, $\phi_1, \phi_2 \geq 0$.
x	The Euclidean distance of two locations, $x \geq 0$.
model	The semivariogram model, it has three options, "Exponential", "Power", "Gaussian" (default:"Exponential").

Value

A scalar representing the spatial isotropic variogram estimate for x.

References

Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu.(2020) Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

GenerateData *Generate three dimensional geographical data.*

Description

The spatial random process $Y(X)$ follows a Gaussian process. The locations coordinates are simulated on a three-dimensional half-ellipsoid centered at (0,0,0) and the projection of the first two dimensions is a disk centered at the origin.

Usage

```
GenerateData(n, SimRep, trueTheta)
```

Arguments

n	The number of locations.
SimRep	The number of realizations of the Gaussian process $Y(X)$.
trueTheta	A given positive value that controls the correlation of the spatial random process.

Value

- LocIndex: three dimensional location coordinates;
- YObs: realizations of the Gaussian process $Y(X)$;
- SigmaObs: covariance matrix of the random process $Y(X)$.

References

- Bornn, L.,Shaddick, G., and Zidek, J.V.(2012). Modeling non-stationary processes through dimension expansion. Journal of American Statistical Association, 107:281–289.
- Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu.(2020). Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

ObjectiveFunBorrn *The objective function in Bornn's paper.*

Description

The OLS fitting procedure is to minimize the objective function,

$$f_{OLS} = \sum_{j<i} \{\hat{\gamma}_{i,j} - \gamma_\phi d_{i,j}([X, Z])\}^2 + \lambda \sum_{k=1}^{p} \|Z_{.k}\|_1$$

Usage

```
ObjectiveFunBorrn(X, para, lambda, Maxd, EmpSemivariogram, model)
```

Arguments

X	The records of location coordinates, a $n \times d$ matrix, where n is the number of locations, and d is the dimension of the record location. Usually, $d = 2$ (latitude and longitude coordinates).
para	A vector-valued parameters (ϕ, Z) in the objective function, dimension (Maxd-2)n + 2.
lambda	A prespecified tuning parameter, $\lambda > 0$.
Maxd	A prespecified maximum coordinate dimension of each location.
EmpSemivariogram	The empirical semivariogram that is computed by EmpSemivariogramFit.
model	The semivariogram model, it has three options, "Exponential", "Power","Gaussian" (default:"Exponential").

Value

An objective value

References

Bornn, L.,Shaddick, G., and Zidek, J.V.(2012). Modeling non-stationary processes through dimension expansion. Journal of American Statistical Association, 107:281–289.

ObjectiveGLSFun *Generalized Least-squares (GLS) objective function.*

Description

The GLS fitting procedure is to minimize the objective function,

$$f_{GLS} = \{vec(U^T) - \gamma_\phi W([X, Z])\}^T \hat{\Sigma}^{-1} \{vec(U^T) - \gamma_\phi W([X, Z])\} + \lambda \sum_{k=1}^{p} \|Z_{.k}\|_1$$

Usage

```
ObjectiveGLSFun(X, para, lambda, Maxd, EmpSemivariogram, model)
```

Arguments

X	The records of location coordinates, a $n \times d$ matrix, where n is the number of locations, and d is the dimension of the record location. Usually, $d = 2$ (latitude and longitude coordinates).
para	A vector-valued parameters (ϕ, Z) in the objective function, dimension (Maxd-2)n + 2.
lambda	A prespecified tuning parameter, $\lambda > 0$.
Maxd	A prespecified maximum coordinate dimension of each location.
EmpSemivariogram	The empirical semivariogram that is computed by EmpSemivariogramFit.
model	The semivariogram model, it has three options, "Exponential", "Power","Gaussian" (default:"Exponential").

Value

An objective value

References

Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu.(2020) Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

ObjectiveWLSFun *Weighted Least-squares (WLS) objective function.*

Description

The WLS fitting procedure is to minimize the objective function,

$$f_{WLS} = \sum_{j<i} \frac{1}{\gamma_\phi^2 d_{i,j}([X,Z])} \{\hat{\gamma}_{i,j} - \gamma_\phi d_{i,j}([X,Z])\}^2 + \lambda \sum_{k=1}^{p} \|Z_{\cdot k}\|_1$$

Usage

```
ObjectiveWLSFun(X, para, lambda, Maxd, EmpSemivariogram, model)
```

Arguments

X	The records of location coordinates, a $n \times d$ matrix, where n is the number of locations, and d is the dimension of the record location. Usually, $d = 2$ (latitude and longitude coordinates).
para	A vector-valued parameters (ϕ, Z) in the objective function, dimension (Maxd-2)n + 2.
lambda	A prespecified tuning parameter, $\lambda > 0$.
Maxd	A prespecified maximum coordinate dimension of each location.

`EmpSemivariogram`
: The empirical semivariogram that is computed by `EmpSemivariogramFit`.

`model`
: The semivariogram model, it has three options, "Exponential", "Power","Gaussian" (default:"Exponential").

Value

An objective value

References

Shanshan Qin, Bin Sun, Yuehua Wu, Yuejiao Fu.(2020) Generalized Least-Squares in Dimension Expansion Method for Nonstationary Processes.

Index