Package: ltsk (via r-universe)

September 8, 2024

Type Package

Title Local Time Space Kriging

Version 1.0.9

Date 2022-04-20

Author Naresh Kumar, Dong Liang, Jin Chen, Jun Chen

Maintainer Dong Liang <dliang@umces.edu>

Description Implements local spatial and local spatiotemporal Kriging based on local spatial and local spatiotemporal variograms, respectively. The method is documented in Kumar et al (2013) <https://www.nature.com/articles/jes201352)>.

License GPL-2

Depends parallel,RANN,R (>= 2.10)

Imports fields,gstat,sp

NeedsCompilation yes

Repository https://dliang-cbl.r-universe.dev

RemoteUrl https://github.com/dliang-cbl/ltsk

RemoteRef HEAD

RemoteSha 02e28d6f9c5d29cd96e8b4db460b4332ff721a8e

Contents

`ltsk-package` *Local Time Space Kriging*

Description

ltsk library is a collection of programs for implementing local spatial and local spatiotemporal Kriging. Unlike global Kriging, ltsk subsets the sample around a given location and time where predicted is needed; estimates variogram using the subset of sample data. Product-sum model is implemented and automatically estimated using the data points within the local neighbourhood. A unique advantage of ltsk is that it addresses non-stationarity, which is difficult to handle in large spatiotemporal dataset.

Details

Package:	ltsk
Type:	Package
Version:	1.0
Date:	2014-12-31
License:	GPL2.0

Author(s)

Naresh Kumar (NKumar@med.miami.edu) Dong Liang (dliang@umces.edu) Jun chen (wdidwlia@gmail.com) Jin Chen (jc.chenjin@gmail.com)

References

Haas, Timothy C. "Local prediction of a spatio-temporal process with an application to wet sulfate deposition." Journal of the American Statistical Association 90.432 (1995): 1189-1199.

Iaco, S. De & Myers, D. E. & Posa, D., 2001. "Space-time analysis using a general product-sum model," Statistics & Probability Letters, Elsevier, vol. 52(1), pages 21-28, March.

Kumar, N., et al. (2013). "Satellite-based PM concentrations and their application to COPD in Cleveland, OH." Journal of Exposure Science and Environmental Epidemiology 23(6): 637-646.

Liang, D. and N. Kumar (2013). "Time-space Kriging to address the spatiotemporal misalignment in the large datasets." Atmospheric Environment 72: 60-69.

dnb *Search Neighbours in Time and Space Within Specified Ranges*

Description

A neighbor search implementation based on ANN trees to identify observed data points within a given distance around location and time interval.

Usage

```
dnb(query, obs, th, xcoord='x',ycoord='y',tcoord='t',
 future=TRUE,by=NULL,nbin=NULL,Large=2000,
 cl=NULL,cluster=c(1,3))
```

Arguments

query	a vector; the x, y coordinates and the time stamp of the query point
obs	a matrix; the x, y coordinates and time stamps of the spatiotemporal locations
th	a vector; the distance threshold and time lag
xcoord	a character constant, the field name for x coordinate in both query and obs
ycoord	a character constant, the field name for y coordinate in both query and obs
tcoord	a character constant, the field name for time coordinate in both query and obs
future	logical, whether include observed spatiotemporal points future in time relative to the query spatiotemporal location.
by	a vector of xcoord, ycoord and zcoord resolutions used in neighbor search
nbin	if by is NULL, a vector of bins to divide the xcoord, ycoord and zcoord data range, Default is 200 for space and unique time units.
Large	a numeric constant, upper limit of neighbor points, beyond which subsampling is performance
cl	a **parallel** cluster object (default NULL means single core)
cluster	if cl is a **parallel** object, the time and space domain in terms of th to divide the query and send to nodes (default (2,10))

Details

Implementation involves first calculating the time lags between query point and observed data (with locational coordinates and time); for observed locations within time lag of query, the function calculates the Euclidean distances between query location and all potential neighbors and select those within specified distance threshold.

The future argument can be used to exclude data in the future in neighbor search. This is useful in an extrapolation application.

Value

A list of vectors, row numbers in the observed data matrix, that are within the given distance threshold and time lag of the query location.

Note

For large dataset, use ANN (for spatial kriging) and Range Tree for spatiotemporal Kriging.

Author(s)

Dong Liang (dliang@umces.edu)

See Also

`get.knn` in **FNN**

Examples

```
data(epa_cl)
coords <- c('x','y','t')
ii <- dnb(query[1,coords],obs[,coords],c(0.1,10))
```

---

ltsk *Ordinary Local Time and Space Kriging*

Description

Function implements ordinary time and space kriging for large data sets, with automatic product-sum variogram estimation.

Usage

```
ltsk(query, obs, th, xcoord = "x", ycoord = "y", tcoord = "t", zcoord = "z",
 by=c(0.001,0.001,1),nbin=NULL,byvariog=c(0.5,0.5,5),subset=T,nmax=20,
 vth = NULL, vlen = NULL, llim = c(3, 3),
 verbose = T, Large = 2000, future=T, cl = NULL, cluster=c(2,10))
```

Arguments

`query`	a data frame containing query spatiotemporal locations for which predictions are needed
`obs`	a data frame containing spatiotemporal locations and observed data
`th`	a vector, distance threshold and time lag to define neighbors of a query point
`xcoord`	a character constant, the field name for x coordinate in both `query` and `obs`
`ycoord`	a character constant, the field name for y coordinate in both `query` and `obs`
`tcoord`	a character constant, the field name for time coordinate in both `query` and `obs`

zcoord	a character constant, the field name for data in obs
by	a vector of xcoord, ycoord and zcoord resolutions used in neighbor search
nbin	if by is NULL, a vector of bins to divide the xcoord, ycoord and zcoord data range, Default is 200 for space and unique time units.
byvariog	a vector of xcoord, ycoord and zcoord resolutions for variogram estimation (NULL means trying to calculate variogram for each query point).
subset	for local kriging: whether subset to observations within estimated time and space range.
nmax	for local kriging: the number of nearest observations that should be used for prediction where nearest is defined in terms of semi-variance of local model. By default all used.
vth	thersholds for local spatiotemporal varigoram (default 75% of the max lag difference)
vlen	numbers of bins for local spatiotemporal varigram(default,space 15, temporal for each day)
llim	lower limits for number of regions and intervals with observed data to calculate Kriging (default 3 spatial regions, 3 temporal intervals)
verbose	logical, whether print details information
Large	a numeric constant, upper limit of neighbor points, beyond which subsampling is performance
future	logical, whether including observed points in future relative to query points.
cl	a **parallel** cluster object (default NULL means single core)
cluster	if cl is a **parallel** object, the time and space domain in terms of th to divide the query and send to nodes (default (2,10))

Details

Function implements automatic variogram estimation (when possible) within a local spatiotemporal neighborhoods, and ordinary krigng based on the produce-sum variogram within that neighborhood. An variogram is estimated for each query point to allow for possible non-stationarity in the data generating field.

If the number of neighbors exceeds a user-specified upper limit (Large), neighbors are sub-sampled in a balanced way to reduce the neighborhood size.

Four variogram models: Gaussian, exponential, spherical and Matern are automatically fit to the empirical space and time variogram in the first lag. The range parameter is estimated from the first distance lag where the empirical variogram exceeds 80% of the maximum. Weighted least square is then used to estimate the nugget and partial sill parameters. Model with minimal residual sum of squares between the empirical and fitted variogram is chosen as the variogram model.

Value

Kriging mean and standard deviation and quality flags.

0 valid prediction

1	not enough temporal neighbors
2	not enough spatial neighbors
3	not enough neighbors
4	variogram could not be fit

Author(s)

Naresh Kumar (NKumar@med.miami.edu) Dong Liang (dliang@umces.edu)

References

Haas, Timothy C. "Local prediction of a spatio-temporal process with an application to wet sulfate deposition." Journal of the American Statistical Association 90.432 (1995): 1189-1199.

Iaco, S. De & Myers, D. E. & Posa, D., 2001. "Space-time analysis using a general product-sum model," Statistics & Probability Letters, Elsevier, vol. 52(1), pages 21-28, March.

Kumar, N., et al. (2013). "Satellite-based PM concentrations and their application to COPD in Cleveland, OH." Journal of Exposure Science and Environmental Epidemiology 23(6): 637-646.

Liang, D. and N. Kumar (2013). "Time-space Kriging to address the spatiotemporal misalignment in the large datasets." Atmospheric Environment 72: 60-69.

Examples

```
## load the data
data(ex)
data(epa_cl)
## apply log transformation
obs[,'pr_pm25'] = log(obs[,'pr_pm25'])
## run kriging
system.time(out <- ltsk(ex2.query[1:2,],obs,c(0.10,10),zcoord='pr_pm25',verbose=FALSE))
table(out$flag)
```

---

ltsk-interval *Internal functions to ltsk*

Description

These functions are working R functions that are called by the ltsk function. They should not be directly used.

obs *example data sets for Cleveland OH*

Description

query and observed data for Cleveland OH

Usage

```
data(epa_cl)
```

Index