

Package ‘psgp’

April 12, 2012

Version 0.3-1

Date 2012-01-23

Title Projected Spatial Gaussian Process (psgp) methods

Author Ben Ingram <bri@utalca.cl>, Remi Barillec <r.barillec@aston.ac.uk>

Maintainer Remi Barillec <r.barillec@aston.ac.uk>

Depends intamap, Rcpp (>= 0.8.0), RcppArmadillo (>= 0.2.0), doSNOW

LinkingTo Rcpp, RcppArmadillo

Description Implements projected sparse Gaussian process kriging for the intamap package

License GPL (>= 2)

Repository CRAN

Date/Publication 2012-04-12 18:48:02

R topics documented:

| | |
|-----------------------------------|---|
| buildMetadata | 2 |
| estimateParameters.psgp | 3 |
| learnParameters | 4 |
| makePrediction | 6 |
| psgp | 7 |
| spatialPredict.psgp | 8 |

| | |
|--------------|-----------|
| Index | 10 |
|--------------|-----------|

| | |
|---------------|-----------------------------|
| buildMetadata | <i>Build Metadata Table</i> |
|---------------|-----------------------------|

Description

buildMetadata builds a metadata table of likelihood model descriptors in the PSGP framework. This is an internal function and it should not be used directly.

Usage

```
buildMetadata(observations)
```

Arguments

observations an observations data frame containing a vector of observation variances (observations\$oevar) and, optionally, a list of observation biases (observations\$oebias). If the biases are omitted, a zero bias is assumed for all likelihood models.

Details

buildMetadata builds a metadata table of likelihood model descriptors in the PSGP framework. The likelihood models are assumed Gaussian with variances specified in the vector observations\$oevar (the bias is assumed to be zero). Optionally, biases can be specified in the observations\$oebias vector. However, biases are not taken into account in the current version of the PSGP package (they will be in a future release).

Author(s)

Remi Barillec

References

<http://www.intamap.org/>

See Also

[learnParameters](#), [makePrediction](#), [estimateParameters](#), [spatialPredict](#),

Examples

```
## Load our favourite dataset
data(meuse)
obs <- meuse

## Number of observations
nobs <- length(obs$y)

## Indicate which likelihood model should be used for each observation
obs$oeid <- seq(1:nobs)
```

```
## Use random variances for the sake of the example
obs$oevar <- rnorm( max(obs$oeid) )

## Generate metadata table and print it out
metadata <- buildMetadata(obs)
print(metadata)
```

estimateParameters.psgp

Parameter estimation using a Projected Sequential Gaussian Process (PSGP)

Description

This overloads the [estimateParameters](#) routine from the intamap package for interpolation using the PSGP method.

Usage

```
estimateParameters(object, ...)
```

Arguments

| | |
|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| object | a list object of Intamap type. Most arguments necessary for interpolation are passed through this object. See intamap-package for further description of the necessary content of this variable. |
| ... | other parameters for the generic method, not used for this method |

Details

See [psgp-package](#) and [learnParameters](#) for further details.

Author(s)

Remi Barillec, Ben Ingram

See Also

[learnParameters](#), [estimateParameters](#), [makePrediction](#), [createIntamapObject](#)

Examples

```

# load our favourite dataset
data(meuse)
coordinates(meuse) = ~x+y
meuse$value = log(meuse$zinc)
data(meuse.grid)
gridded(meuse.grid) = ~x+y
proj4string(meuse) = CRS("+init=epsg:28992")
proj4string(meuse.grid) = CRS("+init=epsg:28992")

# the following two steps are only needed if one wishes to
# include observation errors

# indicate which likelihood model should be used for each observation
# in this case we use a different model for each observation
nobs = length(meuse$value)      # Number of observations
meuse$oeid <- seq(1:nobs)

# the variances for the error models are random in this example
# in real examples they will come from actual measurements
# characteristics
meuse$oevar <- abs( rnorm( max(meuse$oeid) ) )

# set up intamap object:
obj = createIntamapObject(
  observations = meuse,
  predictionLocations = meuse.grid,
  targetCRS = "+init=epsg:3035",
  class = "psgp"    # Use PSGP for parameter estimation/interpolation
)

# do interpolation step:
obj = conformProjections(obj)
obj = estimateParameters(obj)

```

learnParameters

Projected Sequential Gaussian Process

Description

learnParameters performs maximum likelihood parameter estimation in the PSGP framework.

Usage

```
learnParameters(object)
```

Usage

```
learnParameters(object)
```

Arguments

`object` a list object of `intamap` type. Most arguments necessary for interpolation are passed through this object. See [intamap-package](#) for further description of the necessary content of this variable.

Details

The Projected Spatial Gaussian Process (PSGP) framework provides an approximation to the full Gaussian process in which the observations are projected sequentially onto an optimal subset of 'active' observations. Spatial interpolation is done using a mixture of covariance kernels (exponential and Matern 5/2).

The function `learnParameters` is an internal function for estimating the parameters of the covariance function given the data, using a maximum likelihood approach. A valid `intamap` object must be passed in.

PSGP is able to also take the measurement characteristics (i.e. errors) into account using possibly many error models. For each error model, assumed Gaussian, the error variance can be specified. The vector `object$observations$oevar` contains all variances for the error models (one value per error model).

Which error model is used for a given observation is determined by the `object$observations$oeid` vector of indices, which specifies the index of the model to be used for each observation.

Warning

It is advised to use the `intamap` wrapper [estimateParameters](#) rather than calling this method directly.

Author(s)

Ben Ingram, Remi Barillec

References

Csato and Opper, 2002; Ingram et al., 2008

<http://www.intamap.org/>

See Also

[makePrediction](#), [learnParameters](#), [estimateParameters](#), [createIntamapObject](#)

Examples

```
# see example in estimateParameters
```

| | |
|----------------|---------------------------------------------------|
| makePrediction | <i>Spatial projected sequential GP prediction</i> |
|----------------|---------------------------------------------------|

Description

makePrediction performs prediction/interpolation within the PSGP package.

Usage

```
makePrediction(object, vario)
```

Usage

```
makePrediction(object, vario)
```

Arguments

| | |
|--------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| object | a list object of intamap type. Most arguments necessary for interpolation are passed through this object. See intamap-package for further description of the necessary content of this variable. Additional meta data about the measurement process is included in this object. In particular, see learnParameters for a way to specify measurement error variances. |
| vario | Log-parameters of the covariance function. For compatibility with the intamap package, the log-parameters of the PSGP covariance function are stored within a variogram array object (see vgm), as follows: vario[1,1] NA vario[1,2] length scale (or range) of the Exponential kernel vario[1,3] variance (or sill) of the Exponential kernel vario[1,4] length scale (or range) of the Matern 5/2 kernel vario[1,5] variance (or sill) of the Matern 5/2 kernel vario[1,6] inverse bias (i.e. 1/mean(data)) vario[1,7] white noise variance (nugget) |

Details

The Projected Spatial Gaussian Process (PSGP) framework provides an approximation to the full Gaussian process in which the observations are projected sequentially onto an optimal subset of 'active' observations. Spatial interpolation is done using a mixture of covariance kernels (Exponential and Matern 5/2).

The function makePrediction is a function for making predictions at a set of unobserved inputs (or locations).

Measurement characteristics (i.e. observation error) can be specified if needed. See [learnParameters](#) for a description of how to specify measurement error models with given variances.

Warning

It is advised to use the intamap wrapper [spatialPredict](#) rather than calling this method directly.

Author(s)

Ben Ingram, Remi Barillec

References

<http://www.intamap.org/>

L. Csato and M. Opper. Sparse online Gaussian processes. *Neural Computation*, 14(3): 641-669, 2002.

B. Ingram, D. Cornford, and D. Evans. Fast algorithms for automatic mapping with space- limited covariance functions. *Stochastic Environmental Research and Risk Assessment*, 22 (5):661-670, 2008.

See Also

[learnParameters](#) [spatialPredict](#), [createIntamapObject](#)

Examples

```
# see example in spatialPredict
```

psgp

Projected Sequential Gaussian Processes for spatial interpolation

Description

The psgp package provides a spatial interpolation method based on Projected Sequential Gaussian Processes (PSGP) for the intamap package. The PSGP is an approximation to the standard Gaussian process whereby the observations are projected onto a subset of optimal "active" observations, thus reducing possible redundancy in the data and allowing for faster, memory efficient, interpolation. The projection is done in a sequential manner, that is one observation is projected onto the active subset at a time. This allows for larger datasets to be processed, and overcomes the limitations of standard Gaussian processes related to storing the full covariance matrix (which can be unfeasible for really large datasets).

This implementation of PSGP for spatial interpolation uses a mixture of covariance kernels, namely an Exponential kernel and a Matern kernel with roughness parameter $5/2$. The covariance function also includes a nugget term (white noise) and bias term.

Parameter estimation

The method [estimateParameters](#) looks for a set of covariance function parameters (kernels, white noise and bias) which maximise the likelihood of the observation.

Prediction/Interpolation

The method [spatialPredict](#) computes predictions (including variance) at a set of unobserved locations.

Measurement errors

It is possible to include measurement errors if these are available. These will be taken into account when estimating parameters and making predictions. See [learnParameters](#) for more information.

System requirements

the PSGP package is written in C++ and uses the Armadillo library for all linear algebra routines.

Author(s)

Ben Ingram, Remi Barillec

References

<http://www.intamap.org>

<http://arma.sourceforge.net>

| | |
|---------------------|--------------------------------------------------------------------------------|
| spatialPredict.psgp | <i>Spatial prediction using a Projected Sequential Gaussian Process (PSGP)</i> |
|---------------------|--------------------------------------------------------------------------------|

Description

This overloads the [spatialPredict](#) routine from the intamap package for interpolation using the PSGP method.

Usage

```
spatialPredict(object, ...)
```

Arguments

`object` a list object of type PSGP. Most arguments necessary for interpolation are passed through this object. See [intamap-package](#) for further description of the necessary content of this variable

`...` optional extra arguments (these are only used for debugging purposes)

Details

See [psgp-package](#) and [makePrediction](#) for further detail.

Author(s)

Ben Ingram, Remi Barillec

See Also

[psgp-package](#), [estimateParameters](#), [makePrediction](#) [createIntamapObject](#)

Examples

```
# Load our favourite data set
data(meuse)
coordinates(meuse) = ~x+y
meuse$value = log(meuse$zinc)
data(meuse.grid)
gridded(meuse.grid) = ~x+y
proj4string(meuse) = CRS("+init=epsg:28992")
proj4string(meuse.grid) = CRS("+init=epsg:28992")

# Specify a different observation error model for each observation
nobs = length(meuse$value)      # Number of observations
meuse$oeid = seq(1:nobs)        # One error model per observation

# Indicate the variance for each of these error models
meuse$oevar <- abs( rnorm( max(meuse$oeid) ) )

# Set up intamap object
obj = createIntamapObject(
  observations = meuse,
  predictionLocations = meuse.grid,
  targetCRS = "+init=epsg:3035",
  class = "psgp"
)

# Estimate parameters and predict at new locations (interpolation)
obj = conformProjections(obj)
obj = estimateParameters(obj)
obj = spatialPredict(obj)

# Plot results
plotIntamap(obj)
```

Index

*Topic **spatial**

- buildMetadata, [2](#)
- estimateParameters.psgp, [3](#)
- learnParameters, [4](#)
- makePrediction, [6](#)
- spatialPredict.psgp, [8](#)

buildMetadata, [2](#)

createIntamapObject, [3](#), [5](#), [7](#), [8](#)

estimateParameters, [2](#), [3](#), [5](#), [7](#), [8](#)

estimateParameters
 (estimateParameters.psgp), [3](#)

estimateParameters.psgp, [3](#)

intamap-package, [3](#), [6](#)

learnParameters, [2](#), [3](#), [4](#), [5–8](#)

makePrediction, [2](#), [3](#), [5](#), [6](#), [8](#)

psgp, [7](#)

psgp-package (psgp), [7](#)

spatialPredict, [2](#), [6–8](#)

spatialPredict (spatialPredict.psgp), [8](#)

spatialPredict.psgp, [8](#)

vgm, [6](#)