Package ‘GPBayes’

August 6, 2022

Type Package

Title Tools for Gaussian Process Modeling in Uncertainty Quantification

Version 0.1.0-4

Date 2022-8-1

Maintainer Pulong Ma <mpulong@gmail.com>

Description Gaussian processes (GPs) have been widely used to model spatial data, spatio-temporal data, and computer experiments in diverse areas of statistics including spatial statistics, spatio-temporal statistics, uncertainty quantification, and machine learning. This package creates basic tools for fitting and prediction based on GPs with spatial data, spatio-temporal data, and computer experiments. Key characteristics for this GP tool include: (1) the comprehensive implementation of various covariance functions including the Matérn family and the Confluent Hypergeometric family with isotropic form, tensor form, and automatic relevance determination form, where the isotropic form is widely used in spatial statistics, the tensor form is widely used in design and analysis of computer experiments and uncertainty quantification, and the automatic relevance determination form is widely used in machine learning; (2) implementations via Markov chain Monte Carlo (MCMC) algorithms and optimization algorithms for GP models with all the implemented covariance functions. The methods for fitting and prediction are mainly implemented in a Bayesian framework; (3) model evaluation via Fisher information and predictive metrics such as predictive scores; (4) built-in functionality for simulating GPs with all the implemented covariance functions; (5) unified implementation to allow easy specification of various GPs.

License GPL (>= 2)

Encoding UTF-8

BugReports https://github.com/pulongma/GPBayes/issues

Imports Rcpp (>= 1.0.1), stats, methods

LinkingTo Rcpp, RcppEigen, RcppProgress

SystemRequirements GNU Scientific Library version >= 2.5

NeedsCompilation yes

Author Pulong Ma [aut, cre]

RoxygenNote 7.1.1

Repository CRAN

Date/Publication 2022-08-06 06:00:02 UTC

R topics documented:

`GPBayes-package` *Tools for Gaussian Stochastic Process Modeling in Uncertainty Quantification*

Description

Gaussian processes (GPs) have been widely used to model spatial data, spatio-temporal data, and computer experiments in diverse areas of statistics including spatial statistics, spatio-temporal statistics, uncertainty quantification, and machine learning. This package creates basic tools for fitting and prediction based on GPs with spatial data, spatio-temporal data, and computer experiments. Key characteristics for this GP tool include: (1) the comprehensive implementation of various covariance functions including the Matérn family and the Confluent Hypergeometric family with isotropic form, tensor form, and automatic relevance determination form, where the isotropic form is widely used in spatial statistics, the tensor form is widely used in design and analysis of computer experiments and uncertainty quantification, and the automatic relevance determination form is widely used in machine learning; (2) implementations via Markov chain Monte Carlo (MCMC) algorithms and optimization algorithms for GP models with all the implemented covariance functions. The methods for fitting and prediction are mainly implemented in a Bayesian framework; (3) model evaluation via Fisher information and predictive metrics such as predictive scores; (4) built-in functionality for simulating GPs with all the implemented covariance functions; (5) unified implementation to allow easy specification of various GPs.

Details

- Data types: For many scientific applications, spatial data, spatio-temporal data, and computer experiments arise naturally. This package provides a comprehensive set of basic tools to fit GaSP models for univariate and multivariate spatial data, spatio-temporal data, computer experiments. Various covariance functions have been implemented including the Confluent Hypergeometric covariance functions, the Matérn covariance functions, the Gaussian covariance function, the generalized Cauchy covariance function. These covariance families can be in isotropic form, in tensor form, or in automatic relevance determination form. The routines `kernel` and `ikernel` contain the details of implementation.
- Model simulation: This package can simulate realizations from GaSP for different types of data including spatial data, spatio-temporal data, and computer experiments. This feature is quite useful in part because benchmarks are used to evaluate the performance of GaSP models. This functionality is implemented in the routine `gp.sim`.
- Model fitting: Both maximum likelihood methods (or its variants) and Bayes estimation methods such as maximum a posterior (MAP) and Markov chain Monte Carlo (MCMC) methods are implemented. In this package, the nugget parameter is included in the model by default for the sake of better prediction performance and stable computation in practice. In addition, the smoothness parameter in covariance functions such as the Matérn class and the Confluent Hypergeometric class can be estimated. The routine `gp.optim` provides optimization based estimation approaches and the routine `gp.mcmc` provides MCMC algorithms based estimation approaches.
- Model prediction: Prediction is made based on the parameter estimation procedure. If maximum likelihood estimation (MLE) methods are used for parameter estimation, the plug-in approach is used for prediction in the sense that MLEs of parameters are plugged into posterior predictive distributions. If partial Bayes methods (e.g., maximum a posterior) are used, the plug-in approach is used for prediction as well. If fully Bayes methods via MCMC algorithms are used, posterior samples are drawn from posterior predictive distributions. The routine `gp.mcmc` allows prediction to be made within the MCMC algorithms, while the routine `gp.predict` generates prediction with estimated parameters.
- Model assessment: Tools for assessing model adequacy are included in a Bayesian context. Deviance information criteria (DIC), log pointwise predictive density, and log joint predictive density can be computed via the routine `gp.model.adequacy`.

Author(s)

Pulong Ma <mpulong@gmail.com>

References

- Cressie, N. (1993). “Statistics for Spatial Data.” John Wiley & Sons, New York, revised edition.

- Ma and Bhadra (2019). "Beyond Matérn: On a Class of Interpretable Confluent Hypergeometric Covariance Functions." arXiv: 1911.05865. https://arxiv.org/abs/1911.05865.
- Sacks, Jerome, William J Welch, Toby J Mitchell, and Henry P Wynn. (1989). "Design and Analysis of Computer Experiments." *Statistical Science* **4**(4). Institute of Mathematical Statistics: 409–435.
- Santner, Thomas J., Brian J. Williams, and William I. Notz. (2018). "The Design and Analysis of Computer Experiments"; 2nd Ed. New York: Springer.
- Stein, Michael L. (1999). "Interpolation of Spatial Data." Springer Science & Business Media, New York.

See Also

GaSP

Examples

```
#####################################################################


#####################################################################
############## Examples for fitting univariate GP models ############

## Set up the Sine example from the tgp package
 code = function(x){
  y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
 }
 n=100
 input = seq(0, 20, length=n)
 XX = seq(0, 20, length=99)
 Ztrue = code(input)
 set.seed(1234)
 output = Ztrue + rnorm(length(Ztrue), sd=0.1)
 df.data = data.frame(x=c(input), y=output, y.true=Ztrue)

 ## fitting a GaSP model with the Cauchy prior
 fit = GaSP(formula=~1, output, input,
           param=list(range=3, nugget=0.1, nu=2.5),
           smooth.est=FALSE, input.new=XX,
           cov.model=list(family="matern", form="isotropic"),
           proposal=list(range=.35, nugget=.8, nu=0.8),
           dtype="Euclidean", model.fit="Cauchy_prior", nsample=3000,
           burnin=500, verbose=TRUE)

 ## fitting a GaSP model with the beta prior
 fit = GaSP(formula=~1, output, input,
           param=list(range=3, nugget=0.1, nu=2.5),
           smooth.est=FALSE, input.new=XX,
           cov.model=list(family="matern", form="isotropic"),
           prior=list(range=list(a=1,b=1,lb=0,ub=20),
                   nugget=list(a=1,b=1,lb=0,ub=var(output)),
           proposal=list(range=.35, nugget=.8, nu=0.8),
```

```
            dtype="Euclidean", model.fit="Beta_prior", nsample=3000,
            burnin=500, verbose=TRUE))

## fitting a GaSP model with the marginal maximum likelihood approach
 fit = GaSP(formula=~1, output, input,
            param=list(range=3, nugget=0.1, nu=2.5),
            smooth.est=FALSE, input.new=XX,
            cov.model=list(family="matern", form="isotropic"),
            dtype="Euclidean", model.fit="MMLE", verbose=TRUE)

## fitting a GaSP model with the profile maximum likelihood approach
 fit = GaSP(formula=~1, output, input,
            param=list(range=3, nugget=0.1, nu=2.5),
            smooth.est=FALSE, input.new=XX,
            cov.model=list(family="matern", form="isotropic"),
            dtype="Euclidean", model.fit="MPLE", verbose=TRUE)
```

BesselK *Modified Bessel function of the second kind*

Description

This function calls the GSL scientific library to evaluate the modified Bessel function of the second kind.

Usage

```
BesselK(nu, z)
```

Arguments

nu	a real positive value
z	a real positive value

Value

a numerical value

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

matern

cauchy *The generalized Cauchy correlation function*

Description

This function computes the generalized Cauchy correlation function given a distance matrix. The generalized Cauchy covariance is given by

$$C(h) = \left\{ 1 + \left(\frac{h}{\phi} \right)^{\nu} \right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter. The case where $\nu = 2$ corresponds to the Cauchy covariance model, which is infinitely differentiable.

Usage

```
cauchy(d, range, tail, nu)
```

Arguments

`d`	a matrix of distances
`range`	a numerical value containing the range parameter
`tail`	a numerical value containing the tail decay parameter
`nu`	a numerical value containing the smoothness parameter

Value

a numerical matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

`kernel`

CH *The Confluent Hypergeometric correlation function proposed by Ma and Bhadra (2019)*

Description

This function computes the Confluent Hypergeometric correlation function given a distance matrix. The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)}\mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

Usage

```
CH(d, range, tail, nu)
```

Arguments

`d`	a matrix of distances
`range`	a numerical value containing the range parameter
`tail`	a numerical value containing the tail decay parameter
`nu`	a numerical value containing the smoothness parameter

Value

a numerical matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, `GaSP`, `gp`, `matern`, `kernel`, `ikernel` cor.to.par *Find the correlation parameter given effective range*

Description

This function finds the correlation parameter given effective range

Usage

```
cor.to.par(
  d,
  param,
  family = "CH",
  cor.target = 0.05,
  lower = NULL,
  upper = NULL,
  tol = .Machine$double.eps
)
```

Arguments

`d` a numerical value containing the effective range

`param` a list containing correlation parameters. The specification of **param** should depend on the covariance model. If the parameter value is `NULL`, this function will find its value given the effective range via root-finding function `uniroot`.

- For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`family` a string indicating the type of covariance structure. The following correlation functions are implemented:

CH The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu + \alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1 - \nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{ 1 + \left(\frac{h}{\phi} \right)^{\nu} \right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi} \right)^{\nu} \mathcal{K}_{\nu} \left(\frac{h}{\phi} \right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

`cor.target` a numerical value. The default value is 0.05, which means that correlation parameters are searched such that the correlation is approximately 0.05.

`lower` a numerical value. This sets the lower bound to find the correlation parameter via the `R` function `uniroot`.

`upper` a numerical value. This sets the upper bound to find the correlation parameter via the `R` function `uniroot`.

`tol` a numerical value. This sets the precision of the solution with default value specified as the machine precision `.Machine$double.eps` in `R`.

Value

a numerical value of correlation parameters

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, `GaSP`, `kernel`, `ikernel`

Examples

```
range = cor.to.par(1,param=list(tail=0.5,nu=2.5), family="CH")
tail = cor.to.par(1,param=list(range=0.5,nu=2.5), family="CH")
range = cor.to.par(1,param=list(nu=2.5),family="matern")
```

`deriv_kernel` *A wraper to construct the derivative of correlation matrix with respect to correlation parameters*

Description

This function wraps existing built-in routines to construct the derivative of correlation matrix with respect to correlation parameters.

Usage

```
deriv_kernel(d, range, tail, nu, covmodel)
```

Arguments

- `d` a matrix or a list of distances returned from `distance`.
- `range` a vector of range parameters
- `tail` a vector of tail decay parameters
- `nu` a vector of smoothness parameters
- `covmodel` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

  family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)}\mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

  where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

  cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

  where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

  matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{h}{\phi}\right)^{\nu}\mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

  where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

- **exp** This is the Matérn correlation with $\nu = 0.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 0.5$.
- **matern_3_2** This is the Matérn correlation with $\nu = 1.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 1.5$.
- **matern_5_2** This is the Matérn correlation with $\nu = 2.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d}\frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

Value

a list of matrices

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

CH, matern, kernel, GPBayes-package, GaSP

Examples

```
input = seq(0,1,length=10)
d = distance(input,input,type="isotropic",dtype="Euclidean")
dR = deriv_kernel(d,range=0.5,tail=0.2,nu=2.5,
         covmodel=list(family="CH",form="isotropic"))
```

---

distance — *Compute distances for two sets of inputs*

Description

This function computes distances for two sets of inputs and returns a R object.

Usage

```
distance(input1, input2, type = "isotropic", dtype = "Euclidean")
```

Arguments

input1	a matrix of inputs
input2	a matrix of inputs
type	a string indicating the form of distances with three froms supported currently: **isotropic**, **tensor**, **ARD**.
dtype	a string indicating distance type: **Euclidean**, **GCD**, where the latter indicates great circle distance.

Value

a R object holding distances for two sets of inputs. If **type** is **isotropic**, a matrix of distances is returned; if **type** is **tensor** or **ARD**, a list of distance matrices along each input dimension is returned.

a numeric vector or matrix of distances

Author(s)

Pulong Ma <mpulong@gmail.com>

Examples

```
input = seq(0,1,length=20)
d = distance(input, input, type="isotropic", dtype="Euclidean")
```

GaSP *Building, fitting, predicting for a GaSP model*

Description

This function serves as a wrapper to build, fit, and make prediction for a Gaussian process model. It calls on functions gp, gp.mcmc, gp.optim, gp.predict.

Usage

```
GaSP(
  formula = ~1,
  output,
  input,
  param,
  smooth.est = FALSE,
  input.new = NULL,
  cov.model = list(family = "CH", form = "isotropic"),
  model.fit = "Cauchy_prior",
  prior = list(),
  proposal = list(range = 0.35, tail = 2, nugget = 0.8, nu = 0.8),
  nsample = 5000,
  burnin = 1000,
  opt = NULL,
  bound = NULL,
  dtype = "Euclidean",
  verbose = TRUE
)
```

Arguments

`formula`	an object of `formula` class that specifies regressors; see formula for details.
`output`	a numerical vector including observations or outputs in a GaSP
`input`	a matrix including inputs in a GaSP
`param`	a list including values for regression parameters, covariance parameters, and nugget variance parameter. The specification of **param** should depend on the covariance model. • The regression parameters are denoted by **coeff**. Default value is **0**. • The marginal variance or partial sill is denoted by **sig2**. Default value is 1. • The nugget variance parameter is denoted by **nugget** for all covariance models. Default value is 0. • For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.

- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`smooth.est` a logical value indicating whether smoothness parameter will be estimated.

`input.new` a matrix of new input locations

`cov.model` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)}\mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)}\left(\frac{h}{\phi}\right)^{\nu}\mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d}\frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`model.fit` a string indicating the choice of priors on correlation parameters:

Cauchy_prior This indicates that a fully Bayesian approach with objective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with half-Cauchy priors (default).

Ref_prior This indicates that a fully Bayesian approach with objective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with reference priors. This is only supported for isotropic covariance functions. For details, see `gp.mcmc`.

Beta_prior This indicates that a fully Bayesian approach with subjective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with beta priors parameterized as $Beta(a, b, lb, ub)$. In the beta distribution, **lb** and **ub** are the support for correlation parameters, and they should be determined based on domain knowledge. **a** and **b** are two shape parameters with default values at 1, corresponding to the uniform prior over the support (lb, ub).

MPLE This indicates that the *maximum profile likelihood estimation* (**MPLE**) is used.

MMLE This indicates that the *maximum marginal likelihood estimation* (**MMLE**) is used.

MAP This indicates that the marginal/integrated posterior is maximized.

`prior` a list containing tuning parameters in prior distribution. This is used only if a subjective Bayes estimation method with informative priors is used.

`proposal` a list containing tuning parameters in proposal distribution. This is used only if a Bayes estimation method is used.

`nsample` an integer indicating the number of MCMC samples.

`burnin` an integer indicating the burn-in period.

`opt` a list of arguments to setup the `optim` routine. Current implementation uses three arguments:

method The optimization method: `Nelder-Mead` or `L-BFGS-B`.

lower The lower bound for parameters.

upper The upper bound for parameters.

`bound` Default value is `NULL`. Otherwise, it should be a list containing the following elements depending on the covariance class:

nugget a list of bounds for the nugget parameter. It is a list containing lower bound **lb** and upper bound **ub** with default value `list(lb=0, ub=Inf)`.

range a list of bounds for the range parameter. It has default value `range=list(lb=0, ub=Inf)` for the Confluent Hypergeometric covariance, the Matérn covariance, exponential covariance, Gaussian covariance, powered-exponential covariance, and Cauchy covariance. The log of range parameterization is used: $\log(\phi)$.

tail a list of bounds for the tail decay parameter. It has default value `list(lb=0, ub=Inf)`for the Confluent Hypergeometric covariance and the Cauchy covariance.

nu a list of bounds for the smoothness parameter. It has default value `list(lb=0, ub=Inf)` for the Confluent Hypergeometric covariance and the Matérn covariance. when the powered-exponential or Cauchy class is used, it has default value **nu**=`list(lb=0, ub=2)`. This can be achieved by specifying the **lower** bound in `opt`.

`dtype` a string indicating the type of distance:

Euclidean Euclidean distance is used. This is the default choice.

GCD Great circle distance is used for data on sphere.

`verbose` a logical value. If it is `TRUE`, the MCMC progress bar is shown.

Value

a list containing the `S4` object gp and prediction results

Author(s)

Pulong Ma `<mpulong@gmail.com>`

See Also

GPBayes-package, gp, gp.mcmc, gp.optim, gp.predict

Examples

```
code = function(x){
y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
return(y)
}
n=100
input = seq(0, 20, length=n)
XX = seq(0, 20, length=99)
Ztrue = code(input)
set.seed(1234)
output = Ztrue + rnorm(length(Ztrue), sd=0.1)

# fitting a GaSP model with the objective Bayes approach
fit = GaSP(formula=~1, output, input,
          param=list(range=3, nugget=0.1, nu=2.5),
          smooth.est=FALSE, input.new=XX,
          cov.model=list(family="matern", form="isotropic"),
          proposal=list(range=.35, nugget=.8, nu=0.8),
          dtype="Euclidean", model.fit="Cauchy_prior", nsample=50,
          burnin=10, verbose=TRUE)
```

---

gp — *Construct the* S4 *object* gp

Description

This function constructs the S4 object gp that is used for Gaussian process model fitting and prediction.

Usage

```
gp(
  formula = ~1,
  output,
  input,
  param,
  smooth.est = FALSE,
  cov.model = list(family = "CH", form = "isotropic"),
  dtype = "Euclidean"
)
```

Arguments

`formula` an object of `formula` class that specifies regressors; see `formula` for details.

`output` a numerical vector including observations or outputs in a GaSP

`input` a matrix including inputs in a GaSP

`param` a list including values for regression parameters, covariance parameters, and nugget variance parameter. The specification of **param** should depend on the covariance model.

- The regression parameters are denoted by **coeff**. Default value is **0**.
- The marginal variance or partial sill is denoted by **sig2**. Default value is 1.
- The nugget variance parameter is denoted by **nugget** for all covariance models. Default value is 0.
- For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`smooth.est` a logical value indicating whether smoothness parameter will be estimated.

`cov.model` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`dtype` a string indicating the type of distance:

Euclidean Euclidean distance is used. This is the default choice.

GCD Great circle distance is used for data on sphere.

Value

an S4 object of gp class

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP

Examples

```
code = function(x){
y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
return(y)
}
n=100
input = seq(0, 20, length=n)
XX = seq(0, 20, length=99)
Ztrue = code(input)
set.seed(1234)
output = Ztrue + rnorm(length(Ztrue), sd=0.1)
obj = gp(formula=~1, output, input,
        param=list(range=4, nugget=0.1,nu=2.5),
        smooth.est=FALSE,
        cov.model=list(family="matern", form="isotropic"))
```

---

gp-class *The gp class*

Description

This is an S4 class definition for gp in the GaSP package.

Slots

`formula` an object of `formula` class that specifies regressors; see `formula` for details.

`output` a numerical vector including observations or outputs in a GaSP

`input` a matrix including inputs in a GaSP

`param` a list including values for regression parameters, correlation parameters, and nugget variance parameter. The specification of **param** should depend on the covariance model.

- The regression parameters are denoted by **coeff**. Default value is $\mathbf{0}$.
- The marginal variance or partial sill is denoted by **sig2**. Default value is 1.

- The nugget variance parameter is denoted by **nugget** for all covariance models. Default value is 0.
- For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`cov.model` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{ -\left(\frac{h}{\phi}\right)^{\nu} \right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2} \right),$$

where ϕ is the range parameter.

form isotropic This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0 \left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta} \right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`smooth.est` a logical value. If it is `TRUE`, the smoothness parameter will be estimated; otherwise the smoothness is not estimated.

`dtype` a string indicating the type of distance:

Euclidean Euclidean distance is used. This is the default choice.

GCD Great circle distance is used for data on sphere.

`loglik` a numerical value containing the log-likelihood with current `gp` object.

`mcmc` a list containing MCMC samples if available.

`prior` a list containing tuning parameters in prior distribution. This is used only if a Bayes estimation method with informative priors is used.

`proposal` a list containing tuning parameters in proposal distribution. This is used only if a Bayes estimation method is used.

`info` a list containing the maximum distance in the input space. It should be a vector if **isotropic** covariance is used, otherwise it is vector of maximum distances along each input dimension

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP

---

gp.fisher *Fisher information matrix*

Description

This function computes the Fisher information matrix $I(\sigma^2, \boldsymbol{\theta})$ for a Gaussian process model. The standard likelihood is defined as

$$L(\mathbf{b}, \sigma^2, \boldsymbol{\theta}; \mathbf{y}) = \mathcal{N}_n(\mathbf{H}\mathbf{b}, \sigma^2\mathbf{R}),$$

where $\mathbf{y} := (y(\mathbf{x}_1), \ldots, y(\mathbf{x}_n))^\top$ is a vector of n observations. $\mathbf{H}$ is a matrix of covariates, $\mathbf{b}$ is a vector of regression coefficients, σ^2 is the variance parameter, $\boldsymbol{\theta}$ contains correlation parameters and nugget parameter, $\mathbf{R}$ denotes the correlation matrix plus nugget variance on the main diagonal.

The integrated likelihood is defined as

$$L^I(\sigma^2, \boldsymbol{\theta}; \mathbf{y}) = \int L(\mathbf{b}, \sigma^2, \boldsymbol{\theta}; \mathbf{y}) \pi^R(\mathbf{b} \mid \sigma^2, \boldsymbol{\theta}) d\mathbf{b},$$

where $\pi^R(\mathbf{b} \mid \sigma^2, \boldsymbol{\theta}) = 1$ is the conditional Jeffreys-rule (or reference prior) in the model with the above standard likelihood when $(\sigma^2, \boldsymbol{\theta})$ is assumed to be known.

- For the Matérn class, current implementation only computes Fisher information matrix for variance parameter σ^2, range parameter ϕ, and nugget variance parameter τ^2. That is, $I(\sigma^2, \boldsymbol{\theta}) = I(\sigma^2, \phi, \tau^2)$.
- For the Confluent Hypergeometric class, current implementation computes Fisher information matrix for variance parameter σ^2, range parameter β, tail decay parameter α, smoothness parameter ν and nugget variance parameter τ^2. That is, $I(\sigma^2, \boldsymbol{\theta}) = I(\sigma^2, \beta, \alpha, \nu, \tau^2)$.

Usage

```
gp.fisher(
  obj = NULL,
  intloglik = FALSE,
  formula = ~1,
  input = NULL,
  param = NULL,
  cov.model = NULL,
  dtype = "Euclidean"
)
```

Arguments

`obj` a gp object. It is optional with default value `NULL`.

`intloglik` a logical value with default value `FALSE`. If it is `FALSE`, Fisher information matrix $I(\sigma^2, \boldsymbol{\theta})$ is derived based on the standard likelihood; otherwise, Fisher information matrix $I(\sigma^2, \boldsymbol{\theta})$ is derived based on the integrated likelihood.

`formula` an object of `formula` class that specifies regressors; see `formula` for details.

`input` a matrix including inputs in a GaSP

`param` a list including values for regression parameters, covariance parameters, and nugget variance parameter. The specification of **param** should depend on the covariance model.

- The regression parameters are denoted by **coeff**. Default value is $\mathbf{0}$.
- The marginal variance or partial sill is denoted by **sig2**. Default value is 1.
- The nugget variance parameter is denoted by **nugget** for all covariance models. Default value is 0.
- For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`cov.model` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`dtype` a string indicating the type of distance:

Euclidean Euclidean distance is used. This is the default choice.

GCD Great circle distance is used for data on sphere.

Value

a numerical matrix of Fisher information

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP, gp, kernel, ikernel,

Examples

```
n=100
input = seq(0, 20, length=n)
range = 1
tail = .5
nu = 1.5
sig2 = 1
nugget = 0.01
coeff = 0
par = list(range=range, tail=tail, nu=nu, sig2=sig2, nugget=nugget, coeff=coeff)
I = gp.fisher(formula=~1, input=input,
        param=list(range=4, nugget=0.1,nu=2.5),
        cov.model=list(family="CH", form="isotropic"))
```

---

gp.get.mcmc *get posterior summary for MCMC samples*

Description

This function processes posterior samples in the gp object.

Usage

```
gp.get.mcmc(obj, burnin = 500)
```

Arguments

obj	a gp object
burnin	a numerical value specifying the burn-in period for calculating posterior summaries.

Value

a list of posterior summaries

See Also

GPBayes-package, `GaSP`, `gp`, `gp.mcmc`

`gp.mcmc` *A wraper to fit a Gaussian stochastic process model with MCMC algorithms*

Description

This function is a wraper to estimate parameters via MCMC algorithms in the GaSP model with different choices of priors.

Usage

```
gp.mcmc(
  obj,
  input.new = NULL,
  method = "Cauchy_prior",
  prior = list(),
  proposal = list(),
  nsample = 10000,
  verbose = TRUE
)
```

Arguments

`obj`	an S4 object gp
`input.new`	a matrix of prediction locations. Default value is `NULL`, indicating that prediction is not carried out along with parameter estimation in the MCMC algorithm.
`method`	a string indicating the Bayes estimation approaches with different choices of priors on correlation parameters: **Cauchy_prior** This indicates that a fully Bayesian approach with objective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with half-Cauchy priors (default). If the smoothness parameter is estimated for `isotropic` covariance functions, the smoothness parameter is assigned with a uniform prior on (0, 4), indicating that the corresponding GP is at most four times mean-square differentiable. This is a reasonable prior belief for modeling spatial processes; If the smoothness parameter is estimated for `tensor` or `ARD` covariance functions, the smoothness parameter is assigned with a uniform prior on (0, 6). **Ref_prior** This indicates that a fully Bayesian approach with objective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with reference priors. If the smoothness parameter is estimated for `isotropic`

covariance functions, the smoothness parameter is assigned with a uniform prior on (0, 4), indicating that the corresponding GP is at most four times mean-square differentiable. This is a reasonable prior belief for modeling spatial processes; If the smoothness parameter is estimated for `tensor` or `ARD` covariance functions, the smoothness parameter is assigned with a uniform prior on (0, 6).

Beta_prior This indicates that a fully Bayesian approach with subjective priors is used for parameter estimation, where location-scale parameters are assigned with constant priors and correlation parameters are assigned with beta priors parameterized as $Beta(a, b, lb, ub)$. In the beta distribution, **lb** and **ub** are the support for correlation parameters, and they should be determined based on domain knowledge. **a** and **b** are two shape parameters with default values at 1, corresponding to the uniform prior over the support (lb, ub).

`prior`	a list containing tuning parameters in prior distributions. This is used only if a Bayes estimation method with subjective priors is used.
`proposal`	a list containing tuning parameters in proposal distributions. This is used only if a Bayes estimation method is used.
`nsample`	an integer indicating the number of MCMC samples.
`verbose`	a logical value. If it is `TRUE`, the MCMC progress bar is shown.

Value

a `gp` object with prior, proposal, MCMC samples included.

Author(s)

Pulong Ma `<mpulong@gmail.com>`

See Also

GPBayes-package, `GaSP`, `gp`, `gp.optim`

Examples

```
code = function(x){
y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
return(y)
}
n=100
input = seq(0, 20, length=n)
XX = seq(0, 20, length=99)
Ztrue = code(input)
set.seed(1234)
output = Ztrue + rnorm(length(Ztrue), sd=0.1)
obj = gp(formula=~1, output, input,
        param=list(range=4, nugget=0.1,nu=2.5),
```

```
        smooth.est=FALSE,
        cov.model=list(family="matern", form="isotropic"))

fit.mcmc = gp.mcmc(obj, method="Cauchy_prior",
                   proposal=list(range=0.3, nugget=0.8),
                   nsample=100, verbose=TRUE)
```

`gp.model.adequacy` *Model assessment based on Deviance information criterion (DIC), logarithmic pointwise predictive density (lppd), and logarithmic joint predictive density (ljpd).*

Description

This function computes effective number of parameters (pD), deviance information criterion (DIC), logarithmic pointwise predictive density (lppd), and logarithmic joint predictive density (ljpd). For detailed introduction of these metrics, see Chapter 7 of Gelman et al. (2013).

The deviance function for a model with a vector of parameters $\boldsymbol{\theta}$ is defined as

$$D(\boldsymbol{\theta}) = -2 \log p(\mathbf{y} \mid \boldsymbol{\theta}),$$

where $\mathbf{y} := (y(\mathbf{x}_1), \dots, y(\mathbf{x}_n))^\top$ is a vector of n observations.

- The effective number of parameters (see p.172 of Gelman et al. 2013) is defined as

  $$pD = E_{\boldsymbol{\theta}|\mathbf{y}}[D(\boldsymbol{\theta})] - D(\hat{\boldsymbol{\theta}}),$$

  where $\hat{\boldsymbol{\theta}} = E_{\boldsymbol{\theta}|\mathbf{y}}[\boldsymbol{\theta}]$. The interpretation is that the effective number of parameters is the "expected" deviance minus the "fitted" deviance. Higher pD implies more over-fitting with estimate $\hat{\boldsymbol{\theta}}$.
- The Deviance information criteria (DIC) (see pp. 172-173 of Gelman et al. 2013) is

  $$DIC = E_{\boldsymbol{\theta}|\mathbf{y}}[D(\boldsymbol{\theta})] + pD.$$

  DIC approximates Akaike information criterion (AIC) and is more appropriate for hierarchical models than AIC and BIC.
- The log predictive density (**lpd**) is defined as

  $$p(y(\mathbf{x}_0) \mid \mathbf{y}) = \int p(y(\mathbf{x}_0) \mid \boldsymbol{\theta}, \mathbf{y}) p(\boldsymbol{\theta} \mid \mathbf{y}) d\boldsymbol{\theta},$$

  where $\mathbf{y} := (y(\mathbf{x}_1), \dots, y(\mathbf{x}_n))^\top$ is a vector of n observations. $\boldsymbol{\theta}$ contains correlation parameters and nugget parameter. This predictive density should be understood as an update of the likelihood since data is treated as prior information now. With a set of prediction locations $\mathcal{X} := \{x_0^i : i = 1, \dots, m\}$. The log pointwise predictive density (**lppd**) is defined as

  $$lppd = \sum_{i=1}^{m} \log p(y(\mathbf{x}_0^i) \mid \mathbf{y}).$$

The log joint predictive density (**ljpd**) is defined as

$$ljpd = \log p(y(\mathcal{X})).$$

The lppd is connected to cross-validation, while the ljpd measures joint uncertainty across prediction locations.

Usage

```
gp.model.adequacy(
  obj,
  testing.input,
  testing.output,
  pointwise = TRUE,
  joint = TRUE
)
```

Arguments

obj	a gp object.
testing.input	a matrix of testing inputs
testing.output	a vector of testing outputs
pointwise	a logical value with default value TRUE. If it is TRUE, **lppd** is calculated.
joint	a logical value with default value TRUE. If it is TRUE, **ljpd** is calculated.

Value

a list containingg **pD**, **DIC**, **lppd**, **ljpd**.

Author(s)

Pulong Ma <mpulong@gmail.com>

References

- Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin (2013). Bayesian Data Analysis, Third Edition. CRC Press.

See Also

GPBayes-package, GaSP, gp,

`gp.optim` *A wraper to fit a Gaussian stochastic process model with optimization methods*

Description

This function is a wraper to estimate parameters in the GaSP model with different choices of estimation methods using numerical optimization methods.

Usage

```
gp.optim(obj, method = "MMLE", opt = NULL, bound = NULL)
```

Arguments

`obj` an `S4` object gp

`method` a string indicating the parameter estimation method:

- **MPLE** This indicates that the *maximum profile likelihood estimation* (**MPLE**) is used.
- **MMLE** This indicates that the *maximum marginal likelihood estimation* (**MMLE**) is used.
- **MAP** This indicates that the marginal/integrated posterior is maximized.

`opt` a list of arguments to setup the `optim` routine. Current implementation uses three arguments:

- **method** The optimization method: `Nelder-Mead` or `L-BFGS-B`.
- **lower** The lower bound for parameters.
- **upper** The upper bound for parameters.

`bound` Default value is `NULL`. Otherwise, it should be a list containing the following elements depending on the covariance class:

- **nugget** a list of bounds for the nugget parameter. It is a list containing lower bound **lb** and upper bound **ub** with default value `list(lb=0, ub=Inf)`.
- **range** a list of bounds for the range parameter. Tt has default value `range=list(lb=0, ub=Inf)` for the Confluent Hypergeometric covariance, the Matérn covariance, exponential covariance, Gaussian covariance, powered-exponential covariance, and Cauchy covariance. The log of range parameterization is used: $\log(\phi)$.
- **tail** a list of bounds for the tail decay parameter. It has default value `list(lb=0, ub=Inf)`for the Confluent Hypergeometric covariance and the Cauchy covariance.
- **nu** a list of bounds for the smoothness parameter. It has default value `list(lb=0, ub=Inf)` for the Confluent Hypergeometric covariance and the Matérn covariance. when the powered-exponential or Cauchy class is used, it has default value **nu**=`list(lb=0, ub=2)`. This can be achived by specifying the **lower** bound in `opt`.

Value

a list of updated gp object **obj** and fitted information **fit**

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP, gp, gp.mcmc

Examples

```
code = function(x){
y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
return(y)
}
n=100
input = seq(0, 20, length=n)
XX = seq(0, 20, length=99)
Ztrue = code(input)
set.seed(1234)
output = Ztrue + rnorm(length(Ztrue), sd=0.1)
obj = gp(formula=~1, output, input,
        param=list(range=4, nugget=0.1,nu=2.5),
        smooth.est=FALSE,
        cov.model=list(family="matern", form="isotropic"))

fit.optim = gp.optim(obj, method="MPLE")
```

---

gp.predict — *Prediction at new inputs based on a Gaussian stochastic process model*

Description

This function provides the capability to make prediction based on a GaSP when different estimation methods are employed.

Usage

```
gp.predict(obj, input.new, method = "Bayes")
```

Arguments

obj an S4 object gp

input.new a matrix of new input lomessageions

method a string indicating the parameter estimation method:

MPLE This indicates that the *maximum profile likelihood estimation* (**MPLE**) is used. This correponds to simple kriging formulas

MMLE This indicates that the *maximum marginal likelihood estimation* (**MMLE**) is used. This corresponds to universal kriging formulas when the vairance parameter is not integrated out. If the variance parameter is integrated out, the predictive variance differs from the universal kriging variance by the factor $\frac{n-q}{n-q-2}$, since the predictive distribution is a Student's t-distribution with degrees of freedom $n-q$.

MAP This indicates that the posterior estimates of model parameters are plugged into the posterior predictive distribution. Thus this approach does not take account into uncertainty in model parameters (**range**, **tail**, **nu**, **nugget**).

Bayes This indicates that a fully Bayesian approach is used for parameter estimation (and hence prediction). This approach takes into account uncertainty in all model parameters.

Value

a list of predictive mean, predictive standard deviation, 95

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP, gp, gp.mcmc, gp.optim

Examples

```
code = function(x){
y = (sin(pi*x/5) + 0.2*cos(4*pi*x/5))*(x<=9.6) + (x/10-1)*(x>9.6)
return(y)
}
n=100
input = seq(0, 20, length=n)
XX = seq(0, 20, length=99)
Ztrue = code(input)
set.seed(1234)
output = Ztrue + rnorm(length(Ztrue), sd=0.1)
obj = gp(formula=~1, output, input,
        param=list(range=4, nugget=0.1,nu=2.5),
        smooth.est=FALSE,
        cov.model=list(family="matern", form="isotropic"))
```

```
fit.optim = gp.optim(obj, method="MMLE")
obj = fit.optim$obj
pred = gp.predict(obj, input.new=XX, method="MMLE")
```

---

gp.sim *Simulate from a Gaussian stochastic process model*

Description

This function simulates realizations from Gaussian processes.

Usage

```
gp.sim(
  formula = ~1,
  input,
  param,
  cov.model = list(family = "CH", form = "isotropic"),
  dtype = "Euclidean",
  nsample = 1,
  seed = NULL
)
```

Arguments

`formula` an object of `formula` class that specifies regressors; see `formula` for details.

`input` a matrix including inputs in a GaSP

`param` a list including values for regression parameters, covariance parameters, and nugget variance parameter. The specification of **param** should depend on the covariance model.

- The regression parameters are denoted by **coeff**. Default value is **0**.
- The marginal variance or partial sill is denoted by **sig2**. Default value is 1.
- The nugget variance parameter is denoted by **nugget** for all covariance models. Default value is 0.
- For the Confluent Hypergeometric class, **range** is used to denote the range parameter β. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.
- For the generalized Cauchy class, **range** is used to denote the range parameter ϕ. **tail** is used to denote the tail decay parameter α. **nu** is used to denote the smoothness parameter ν.

- For the Matérn class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter ν. When $\nu = 0.5$, the Matérn class corresponds to the exponential covariance.
- For the powered-exponential class, **range** is used to denote the range parameter ϕ. **nu** is used to denote the smoothness parameter. When $\nu = 2$, the powered-exponential class corresponds to the Gaussian covariance.

`cov.model` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp The exponential correlation function is given by

$$C(h) = \exp(-h/\phi),$$

where ϕ is the range parameter. This is the Matérn correlation with $\nu = 0.5$.

matern_3_2 The Matérn correlation with $\nu = 1.5$.

matern_5_2 The Matérn correlation with $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`dtype` a string indicating the type of distance:

Euclidean Euclidean distance is used. This is the default choice.

GCD Great circle distance is used for data on sphere.

`nsample` an integer indicating the number of realizations from a Gaussian process

`seed` a number specifying random number seed

Value

a numerical vector or a matrix

Author(s)

Pulong Ma `<mpulong@gmail.com>`

See Also

GPBayes-package, `GaSP`, `gp`

Examples

```
n=50
y.sim = gp.sim(input=seq(0,1,length=n),
               param=list(range=0.5,nugget=0.1,nu=2.5),
               cov.model=list(family="matern",form="isotropic"),
               seed=123)
```

HypergU *Confluent hypergeometric function of the second kind*

Description

This function calls the GSL scientific library to evaluate the confluent hypergeometric function of the second kind; see Abramowitz and Stegun 1972, p.505.

Usage

```
HypergU(a, b, x)
```

Arguments

a	a real value
b	a real value
x	a real value

Value

a numerical value

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

CH

ikernel *A wraper to build different kinds of correlation matrices between two sets of inputs*

Description

This function wraps existing built-in routines to construct a covariance matrix for two input matrices based on data type, covariance type, and distance type. The constructed covariance matrix can be directly used for GaSP fitting and and prediction for spatial data, spatio-temporal data, and computer experiments. This function explicitly takes inputs as arguments. The prefix "i" in ikernel standards for "input".

Usage

```
ikernel(input1, input2, range, tail, nu, covmodel, dtype = "Euclidean")
```

Arguments

input1	a matrix of input locations
input2	a matrix of input locations
range	a vector of range parameters, which could be a scalar.
tail	a vector of tail decay parameters, which could be a scalar.
nu	a vector of smoothness parameters, which could be a scalar.
covmodel	a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp This is the Matérn correlation with $\nu = 0.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 0.5$.

matern_3_2 This is the Matérn correlation with $\nu = 1.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 1.5$.

matern_5_2 This is the Matérn correlation with $\nu = 2.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta}\right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`dtype` a string indicating distance type: **Euclidean**, **GCD**, where the latter indicates great circle distance.

Value

a correlation matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

CH, matern, kernel, GPBayes-package, GaSP

Examples

```
input = seq(0,1,length=10)

cormat = ikernel(input,input,range=0.5,tail=0.2,nu=2.5,
         covmodel=list(family="CH",form="isotropic"))
```

---

kernel *A wraper to build different kinds of correlation matrices with distance as arguments*

Description

This function wraps existing built-in routines to construct a covariance matrix based on data type, covariance type, and distance type with distances as inputs. The constructed covariance matrix can be directly used for GaSP fitting and and prediction for spatial data, spatio-temporal data, and computer experiments.

Usage

```
kernel(d, range, tail, nu, covmodel)
```

Arguments

d	a matrix or a list of distances
range	a vector of range parameters, which could be a scalar.
tail	a vector of tail decay parameters, which could be a scalar.
nu	a vector of smoothness parameters, which could be a scalar.
covmodel	a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family CH The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^\nu\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^\nu \mathcal{K}_\nu\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_\nu(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp This is the Matérn correlation with $\nu = 0.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 0.5$.

matern_3_2 This is the Matérn correlation with $\nu = 1.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 1.5$.

matern_5_2 This is the Matérn correlation with $\nu = 2.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^\nu\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2}\right),$$

where ϕ is the range parameter.

form isotropic This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0 \left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta} \right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

Value

a correlation matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

CH, matern, ikernel, GPBayes-package, GaSP

Examples

```
input = seq(0,1,length=10)
d = distance(input,input,type="isotropic",dtype="Euclidean")
cormat = kernel(d,range=0.5,tail=0.2,nu=2.5,
          covmodel=list(family="CH",form="isotropic"))
```

---

loglik *A wraper to compute the natural logarithm of the integrated likelihood function*

Description

This function wraps existing built-in routines to construct the natural logarithm of the integrated likelihood function. The constructed loglikelihood can be directly used for numerical optimization

Usage

```
loglik(par, output, H, d, covmodel, smooth, smoothness_est)
```

Arguments

`par` a numerical vector, with which numerical optimization routine such as `optim` can be carried out directly. When the confluent Hypergeometric class is used, it is used to hold values for **range**, **tail**, **nugget**, and **nu** if the smoothness parameter is estimated. When the Matérn class or powered-exponential class is used, it is used to hold values for **range**, **nugget**, and **nu** if the smoothness parameter is estimated. The order of the parameter values in `par` cannot be changed. For tensor or ARD form correlation functions, **range** and **tail** becomes a vector.

`output` a matrix of outputs

`H` a matrix of regressors in the mean function of a GaSP model.

`d` an R object holding the distances. It should be a distance matrix for constructing isotropic correlation matrix, or a list of distance matrices along each input dimension for constructing tensor or ARD types of correlation matrix.

`covmodel` a list of two strings: **family**, **form**, where **family** indicates the family of covariance functions including the Confluent Hypergeometric class, the Matérn class, the Cauchy class, the powered-exponential class. **form** indicates the specific form of covariance structures including the isotropic form, tensor form, automatic relevance determination form.

family **CH** The Confluent Hypergeometric correlation function is given by

$$C(h) = \frac{\Gamma(\nu+\alpha)}{\Gamma(\nu)} \mathcal{U}\left(\alpha, 1-\nu, \left(\frac{h}{\beta}\right)^2\right),$$

where α is the tail decay parameter. β is the range parameter. ν is the smoothness parameter. $\mathcal{U}(\cdot)$ is the confluent hypergeometric function of the second kind. For details about this covariance, see Ma and Bhadra (2019) at https://arxiv.org/abs/1911.05865.

cauchy The generalized Cauchy covariance is given by

$$C(h) = \left\{1 + \left(\frac{h}{\phi}\right)^{\nu}\right\}^{-\alpha/\nu},$$

where ϕ is the range parameter. α is the tail decay parameter. ν is the smoothness parameter with default value at 2.

matern The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν.

exp This is the Matérn correlation with $\nu = 0.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 0.5$.

matern_3_2 This is the Matérn correlation with $\nu = 1.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 1.5$.

matern_5_2 This is the Matérn correlation with $\nu = 2.5$. This covariance should be specified as **matern** with smoothness parameter $\nu = 2.5$.

powexp The powered-exponential correlation function is given by

$$C(h) = \exp\left\{ -\left(\frac{h}{\phi}\right)^{\nu} \right\},$$

where ϕ is the range parameter. ν is the smoothness parameter.

gauss The Gaussian correlation function is given by

$$C(h) = \exp\left(-\frac{h^2}{\phi^2} \right),$$

where ϕ is the range parameter.

form **isotropic** This indicates the isotropic form of covariance functions. That is,

$$C(\mathbf{h}) = C^0(\|\mathbf{h}\|; \boldsymbol{\theta}),$$

where $\|\mathbf{h}\|$ denotes the Euclidean distance or the great circle distance for data on sphere. $C^0(\cdot)$ denotes any isotropic covariance family specified in **family**.

tensor This indicates the tensor product of correlation functions. That is,

$$C(\mathbf{h}) = \prod_{i=1}^{d} C^0(|h_i|; \boldsymbol{\theta}_i),$$

where d is the dimension of input space. h_i is the distance along the ith input dimension. This type of covariance structure has been often used in Gaussian process emulation for computer experiments.

ARD This indicates the automatic relevance determination form. That is,

$$C(\mathbf{h}) = C^0\left(\sqrt{\sum_{i=1}^{d} \frac{h_i^2}{\phi_i^2}}; \boldsymbol{\theta} \right),$$

where ϕ_i denotes the range parameter along the ith input dimension.

`smooth` The smoothness parameter ν in a correlation function.

`smoothness_est` a logical value indicating whether the smoothness parameter is estimated.

Value

The natural logarithm of marginal or integrated likelihood

Author(s)

Pulong Ma <`mpulong@gmail.com`>

See Also

`CH`, `matern`, `gp.optim`, GPBayes-package, `GaSP` `matern` *The Matérn correlation function proposed by Matérn (1960)*

Description

This function computes the Matérn correlation function given a distance matrix. The Matérn correlation function is given by

$$C(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{\phi}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{h}{\phi}\right),$$

where ϕ is the range parameter. ν is the smoothness parameter. $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel function of the second kind of order ν. The form of covariance includes the following special cases by specifying ν to be 0.5, 1.5, 2.5.

- $\nu = 0.5$ corresponds to the exponential correlation function (**exp**) of the form

$$C(h) = \exp\left\{-\frac{h}{\phi}\right\}$$

- $\nu = 1.5$ corresponds to the Matérn correlation function with smoothness parameter 1.5 (**matern_3_2**) of the form

$$C(h) = \left(1 + \frac{h}{\phi}\right) \exp\left\{-\frac{h}{\phi}\right\}$$

- $\nu = 2.5$ corresponds to the Matérn correlation function with smoothness parameter 2.5 (**matern_5_2**) of the form

$$C(h) = \left\{1 + \frac{h}{\phi} + \frac{1}{3}\left(\frac{h}{\phi}\right)^2\right\} \exp\left\{-\frac{h}{\phi}\right\}$$

Usage

```
matern(d, range, nu)
```

Arguments

`d`	a matrix of distances
`range`	a numerical value containing the range parameter
`nu`	a numerical value containing the smoothness parameter

Value

a numerical matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

GPBayes-package, GaSP, gp, CH, kernel, ikernel

powexp — *The powered-exponential correlation function*

Description

This function computes the powered-exponential correlation function given a distance matrix. The powered-exponential correlation function is given by

$$C(h) = \exp\left\{-\left(\frac{h}{\phi}\right)^{\nu}\right\},$$

where ϕ is the range parameter. ν is the smoothness parameter. The case $\nu = 2$ corresponds to the well-known Gaussian correlation.

Usage

```
powexp(d, range, nu)
```

Arguments

d	a matrix of distances
range	a numerical value containing the range parameter
nu	a numerical value containing the smoothness parameter

Value

a numerical matrix

Author(s)

Pulong Ma <mpulong@gmail.com>

See Also

kernel

`show,gp-method` *Print the information an object of the gp class*

Description

Print the information an object of the gp class

Usage

```
## S4 method for signature 'gp'
show(object)
```

Arguments

`object`	an object of gp class

Index