Package 'sdPrior'

July 16, 2015

Title Scale-Dependent Hyperpriors in Structured Additive Distributional Regression

Version 0.3

Author Nadja Klein <nklein@uni-goettingen.de>

Maintainer Nadja Klein <nklein@uni-goettingen.de>

Description Utility functions for scale-dependent and alternative hyperpriors.

Depends R (>= 3.1.0)

Imports splines, caTools, GB2, MASS, stats

LazyData true

License GPL-2

NeedsCompilation no

Repository CRAN

Date/Publication 2015-07-16 01:13:54

R topics documented:

---

dapprox_unif *Compute Density Function of Approximated (Differentiably) Uniform Distribution.*

Description

Compute Density Function of Approximated (Differentiably) Uniform Distribution.

Usage

```
dapprox_unif(x, scale, tildec = 13.86294)
```

Arguments

x	denotes the argument of the density function.
scale	the scale parameter originally defining the upper bound of the uniform distribution.
tildec	denotes the ratio between scale parameter θ and s. The latter is responsible for the closeness of the approximation to the uniform distribution. See also below for further details and the default value.

Details

The density of the uniform distribution for τ is approximated by

$$p(\tau) = (1/(1 + exp(\tau\tilde{c}/\theta - \tilde{c})))/(\theta(1 + log(1 + exp(-\tilde{c}))))$$

. This results in

$$p(\tau^2) = 0.5 * (\tau^2)^(-1/2)(1/(1 + exp((\tau^2)^(1/2)\tilde{c}/\theta - \tilde{c})))/(\theta(1 + log(1 + exp(-\tilde{c}))))$$

for tau^2. $\tilde{c}$ is chosen such that $P(\tau <= \theta) >= 0.95$.

Value

the density.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

See Also

rapprox_unif,papprox_unif

DesignM *Computing Designmatrix for Splines*

Description

This function computes the design matrix for Bayesian P-splines as it would be done in BayesX. The implementation currently on works properly for default values (knots=20, degree=3).

Usage

```
DesignM(x, degree = 3, knots = 20, min_x = min(x), max_x = max(x))
```

Arguments

`x`	the covariate vector.
`degree`	of the B-splines, default is 3.
`knots`	number of knots, default is 20.
`min_x`	the left interval boundary, default is min(x).
`max_x`	the right interval boundary, defalut is max(x).

Value

a list with design matrix at distinct covariates, design matrix at all observations, index of sorted observations, the difference matrix, precision matrix and the knots used.

Author(s)

Nadja Klein

References

Stefan Lang and Andy Brezger (2004). Bayesian P-Splines. *Journal of Computational and Graphical Statistics*, **13**, 183–212.

Belitz, C., Brezger, A., Klein, N., Kneib, T., Lang, S., Umlauf, N. (2015): BayesX - Software for Bayesian inference in structured additive regression models. Version 3.0.1. Available from http://www.bayesx.org.

get_theta *Find Scale Parameter for Hyperprior*

Description

This function implements a optimisation routine that computes the scale parameter θ of the scale dependent hyperprior for a given design matrix and prior precision matrix such that approximately $P(|f(x_k| \leq c, k = 1, \ldots, p) \geq 1 - \alpha$

Usage

```
get_theta(alpha = 0.01, method = "integrate", Z, c = 3,
  eps = .Machine$double.eps, Kinv)
```

Arguments

alpha	denotes the 1-α level.
method	either integrate or trapezoid with integrate as default. trapezoid is a self-implemented version of the trapezoid rule.
Z	the design matrix.
c	denotes the expected range of the function.
eps	denotes the error tolerance of the result, default is .Machine$double.eps.
Kinv	the generalised inverse of K.

Value

an object of class list with values from uniroot.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
## Not run:

set.seed(91179)
library(BayesX)
library(MASS)
# prior precision matrix to zambia data set
K <- read.gra(system.file("examples/zambia.gra", package="sdPrior"))
# generalised inverse of K
```

```
Kinv <- ginv(K)

# read data
dat <- read.table(system.file("examples/zambia_height92.raw", package="sdPrior"), header = TRUE)

# design matrix for spatial component
Z <- t(sapply(dat$district, FUN=function(x){1*(x==rownames(K))}))

# get scale parameter
theta <- get_theta(alpha = 0.01, method = "integrate", Z = Z,
                            c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root

## End(Not run)
```

get_theta_aunif — *Find Scale Parameter for Hyperprior for Variances Where the Standard Deviations have an Approximated (Differentiably) Uniform Distribution.*

Description

This function implements a optimisation routine that computes the scale parameter θ of the prior τ^2 (corresponding to a differentiably approximated version of the uniform prior for τ) for a given design matrix and prior precision matrix such that approximately $P(|f(x_k| \leq c, k = 1, \dots, p) \geq 1 - \alpha$

Usage

```
get_theta_aunif(alpha = 0.01, method = "integrate", Z, c = 3,
  eps = .Machine$double.eps, Kinv)
```

Arguments

alpha	denotes the 1-α level.
method	with `integrate` as default. Currently no further method implemented.
Z	the design matrix.
c	denotes the expected range of the function.
eps	denotes the error tolerance of the result, default is `.Machine$double.eps`.
Kinv	the generalised inverse of K.

Value

an object of class `list` with values from `uniroot`.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Andrew Gelman (2006). Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis*, **1**(3), 515–533.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
theta <- get_theta_aunif(alpha = 0.01, method = "integrate", Z = Z,
                             c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root
```

---

get_theta_ga — *Find Scale Parameter for Gamma (Half-Normal) Hyperprior*

Description

This function implements a optimisation routine that computes the scale parameter θ of the gamma prior for τ^2 (corresponding to a half-normal prior for τ) for a given design matrix and prior precision matrix such that approximately $P(|f(x_k| \leq c, k = 1, \ldots, p) \geq 1 - \alpha$

Usage

```
get_theta_ga(alpha = 0.01, method = "integrate", Z, c = 3,
  eps = .Machine$double.eps, Kinv)
```

Arguments

`alpha`	denotes the 1-α level.
`method`	with `integrate` as default. Currently no further method implemented.
`Z`	the design matrix.
`c`	denotes the expected range of the function.
`eps`	denotes the error tolerance of the result, default is `.Machine$double.eps`.
`Kinv`	the generalised inverse of K.

Value

an object of class `list` with values from `uniroot`.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Andrew Gelman (2006). Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis*, **1**(3), 515–533.

Examples

```
set.seed(123)
require(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
theta <- get_theta_ga(alpha = 0.01, method = "integrate", Z = Z,
                            c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root

## Not run:

set.seed(91179)
library(BayesX)
library(MASS)
# prior precision matrix to zambia data set
K <- read.gra(system.file("examples/zambia.gra", package="sdPrior"))
# generalised inverse of K
Kinv <- ginv(K)

# read data
dat <- read.table(system.file("examples/zambia_height92.raw", package="sdPrior"), header = TRUE)

# design matrix for spatial component
Z <- t(sapply(dat$district, FUN=function(x){1*(x==rownames(K))}))

# get scale parameter
theta <- get_theta_ga(alpha = 0.01, method = "integrate", Z = Z,
                            c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root

## End(Not run)
```

`get_theta_gbp` *Find Scale Parameter for Generalised Beta Prime (Half-Cauchy) Hyperprior*

Description

This function implements a optimisation routine that computes the scale parameter θ of the gamma prior for τ^2 (corresponding to a half cauchy for τ) for a given design matrix and prior precision matrix such that approximately $P(|f(x_k| \leq c, k = 1, \ldots, p) \geq 1 - \alpha$

Usage

```
get_theta_gbp(alpha = 0.01, method = "integrate", Z, c = 3,
  eps = .Machine$double.eps, Kinv)
```

Arguments

`alpha`	denotes the 1-α level.
`method`	with `integrate` as default. Currently no further method implemented.
`Z`	the design matrix.
`c`	denotes the expected range of the function.
`eps`	denotes the error tolerance of the result, default is `.Machine$double.eps`.
`Kinv`	the generalised inverse of K.

Value

an object of class `list` with values from `uniroot`.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Andrew Gelman (2006). Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis*, **1**(3), 515–533.

Examples

```
set.seed(123)
require(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
```

```
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
theta <- get_theta_gbp(alpha = 0.01, method = "integrate", Z = Z,
                            c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root

## Not run:

set.seed(91179)
library(BayesX)
library(MASS)
# prior precision matrix to zambia data set
K <- read.gra(system.file("examples/zambia.gra", package="sdPrior"))
# generalised inverse of K
Kinv <- ginv(K)

# read data
dat <- read.table(system.file("examples/zambia_height92.raw", package="sdPrior"), header = TRUE)

# design matrix for spatial component
Z <- t(sapply(dat$district, FUN=function(x){1*(x==rownames(K))}))

# get scale parameter
theta <- get_theta_gbp(alpha = 0.01, method = "integrate", Z = Z,
                            c = 3, eps = .Machine$double.eps, Kinv = Kinv)$root

## End(Not run)
```

---

get_theta_ig *Find Scale Parameter for Inverse Gamma Hyperprior*

Description

This function implements a optimisation routine that computes the scale parameter b of the inverse gamma prior for τ^2 when $a = b = \epsilon$ with ϵ small for a given design matrix and prior precision matrix such that approximately $P(|f(x_k| \leq c, k = 1, \dots, p) \geq 1 - \alpha$ When a unequal to a the shape parameter a has to be specified.

Usage

```
get_theta_ig(alpha = 0.01, method = "integrate", Z, c = 3,
  eps = .Machine$double.eps, Kinv, equals = FALSE, a = 1,
  type = "marginalt")
```

Arguments

alpha	denotes the 1-α level.
method	with `integrate` as default. Currently no further method implemented.
Z	the design matrix.
c	denotes the expected range of the function.
eps	denotes the error tolerance of the result, default is `.Machine$double.eps`.
Kinv	the generalised inverse of K.
equals	saying whether `a=b`. The default is FALSE due to the fact that a is a shape parameter.
a	is the shape parameter of the inverse gamma distribution, default is 1.
type	is either numerical integration (`integrate`) of to obtain the marginal distribution of $z_p'\beta$ or the theoretical marginal t-distribution (`marginalt`). `marginalt` is the default value.

Details

Currently, the implementation only works properly for the cases `a` unequal `b`.

Value

an object of class `list` with values from `uniroot`.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Stefan Lang and Andreas Brezger (2004). Bayesian P-Splines. *Journal of Computational and Graphical Statistics*, **13**, 183-212.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
theta <- get_theta_ig(alpha = 0.01, method = "integrate", Z = Z,
```

```
                         c = 3, eps = .Machine$double.eps, Kinv = Kinv,
     equals = FALSE, a = 1, type="marginalt")$root
```

---

mdf_aunif *Marginal Density for Given Scale Parameter and Approximated Uniform Prior for τ*

Description

This function computes the marginal density of $z_p'\beta$ for approximated uniform hyperprior for τ

Usage

```
mdf_aunif(f, theta, Z, Kinv)
```

Arguments

f	point the marginal density to be evaluated at.
theta	denotes the scale parameter of the approximated uniform hyperprior for τ.
Z	the row of the design matrix evaluated.
Kinv	the generalised inverse of K.

Value

the marginal density evaluated at point x.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
fgrid <- seq(-3,3,length=1000)
mdf <- mdf_aunif(fgrid,theta=0.0028,Z=Z,Kinv=Kinv)
```

mdf_ga *Marginal Density for Given Scale Parameter and Half-Normal Prior for τ*

Description

This function computes the marginal density of $z_p'\beta$ for gamma priors for τ^2 (referring to a half-normal prior for τ).

Usage

```
mdf_ga(f, theta, Z, Kinv)
```

Arguments

f	point the marginal density to be evaluated at.
theta	denotes the scale parameter of the gamma hyperprior for τ^2 (half-normal for τ).
Z	the row of the design matrix evaluated.
Kinv	the generalised inverse of K.

Value

the marginal density evaluated at point x.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
fgrid <- seq(-3,3,length=1000)
mdf <- mdf_ga(fgrid,theta=0.0028,Z=Z,Kinv=Kinv)
```

mdf_gbp	*Marginal Density for Given Scale Parameter and Half-Cauchy Prior for τ*

Description

This function computes the marginal density of $z_p'\beta$ for generalised beta prior hyperprior for τ^2 (half-Chauchy for τ)

Usage

```
mdf_gbp(f, theta, Z, Kinv)
```

Arguments

f	point the marginal density to be evaluated at.
theta	denotes the scale parameter of the generalised beta prior hyperprior for τ^2 (half-Chauchy for τ).
Z	the row of the design matrix evaluated.
Kinv	the generalised inverse of K.

Value

the marginal density evaluated at point x.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
fgrid <- seq(-3,3,length=1000)
mdf <- mdf_gbp(fgrid,theta=0.0028,Z=Z,Kinv=Kinv)
```

mdf_ig *Marginal Density for Given Scale Parameter and Inverse Gamma Prior for τˆ2*

Description

This function computes the marginal density of $z_p'\beta$ for inverse gamma hyperpriors with shape parameter a=1.

Usage

```
mdf_ig(f, theta, Z, Kinv)
```

Arguments

f	point the marginal density to be evaluated at.
theta	denotes the scale parameter of the inverse gamma hyperprior.
Z	the row of the design matrix evaluated.
Kinv	the generalised inverse of K.

Value

the marginal density evaluated at point x.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
fgrid <- seq(-3,3,length=1000)
mdf <- mdf_ig(fgrid,theta=0.0028,Z=Z,Kinv=Kinv)
```

`mdf_sd` *Marginal Density for Given Scale Parameter and Scale-Dependent Prior for τˆ2*

Description

This function computes the marginal density of $z_p'\beta$ for scale-dependent priors for τ^2

Usage

```
mdf_sd(f, theta, Z, Kinv)
```

Arguments

`f`	point the marginal density to be evaluated at.
`theta`	denotes the scale parameter of the scale-dependent hyperprior for τ^2.
`Z`	the row of the design matrix evaluated.
`Kinv`	the generalised inverse of K.

Value

the marginal density evaluated at point x.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

Examples

```
set.seed(123)
library(MASS)
# prior precision matrix (second order differences)
# of a spline of degree l=3 and with m=20 inner knots
# yielding dim(K)=m+l-1=22
K <- t(diff(diag(22), differences=2))%*%diff(diag(22), differences=2)
# generalised inverse of K
Kinv <- ginv(K)
# covariate x
x <- runif(1)
Z <- matrix(DesignM(x)$Z_B,nrow=1)
fgrid <- seq(-3,3,length=1000)
mdf <- mdf_sd(fgrid,theta=0.0028,Z=Z,Kinv=Kinv)
```

papprox_unif *Compute Cumulative Distribution Function of Approximated (Differentiably) Uniform Distribution.*

Description

Compute Cumulative Distribution Function of Approximated (Differentiably) Uniform Distribution.

Usage

```
papprox_unif(x, scale, tildec = 13.86294)
```

Arguments

x	denotes the argument of cumulative distribution function
scale	the scale parameter originally defining the upper bound of the uniform distribution.
tildec	denotes the ratio between scale parameter θ and s. The latter is responsible for the closeness of the approximation to the uniform distribution. See also below for further details and the default value.

Details

The cumulative distribution function of dapprox_unif is given by

$$(1/(log(1 + exp(-\tilde{c})) + \tilde{c})) * (\tilde{c} * (\tau^2)^(1/2)/\theta - log(exp((\tau^2)^(1/2) * \tilde{c}/\theta) + exp(\tilde{c})))$$

$\tilde{c}$ is chosen such that $P(\tau^2 <= \theta) >= 0.95$.

Value

the cumulative distribution function.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

See Also

rapprox_unif,dapprox_unif

rapprox_unif *Draw Random Numbers from Approximated (Differentiably) Uniform Distribution.*

Description

Draw Random Numbers from Approximated (Differentiably) Uniform Distribution.

Usage

```
rapprox_unif(n = 100, scale, tildec = 13.86294, seed = 123)
```

Arguments

n	number of draws.
scale	the scale parameter originally defining the upper bound of the uniform distribution.
tildec	denotes the ratio between scale parameter θ and s. The latter is responsible for the closeness of the approximation to the uniform distribution. See also below for further details and the default value.
seed	denotes the seed

Details

The method is based on the inversion method and the quantile function is computed numerically using uniroot.

Value

n draws with density papprox_unif.

Author(s)

Nadja Klein

References

Nadja Klein and Thomas Kneib (2015). Scale-Dependent Priors for Variance Parameters in Structured Additive Distributional Regression. *Working Paper*.

See Also

rapprox_unif,papprox_unif

`zambia_graph` *Prior precision matrix for spatial variable in Zambia data set*

Description

This is a 57x57 matrix containing row- and columwise the regions of Zambia, and the entries define the neighbourhoodstructure. The corresponding map sambia.bnd can be downloaded from http://www.stat.uni-muenchen.de/~kneib/regressionsbuch/daten_e.html. from the bnd file the prior precision matrix is obtained by library(BayesX) map <- read.bnd("zambia.bnd") K <- bnd2gra(map)

`zambia_height92` *Malnutrition in Zambia*

Description

The primary goal of a statistical analysis is to determine the effect of certain socioeconomic variables of the child, the mother, and the household on the child‘s nutritional condition

- zscore child‘s Z-score
- c _breastf duration of breastfeeding in months
- c_age child‘s age in months
- m_agebirth mother‘s age at birth in years
- m_height mother‘s height in centimeter
- m_bmi mother‘s body mass index
- m_education mother‘s level of education
- m_work mother‘s work status
- region region of residence in Zambia
- district district of residence in Zambia

Format

A data frame with 4421 rows and 21 variables

Source

http://www.stat.uni-muenchen.de/~kneib/regressionsbuch/daten_e.html

Index