

Package ‘GSM’

July 8, 2015

Title Gamma Shape Mixture

Version 1.3.2

Date 2015-07-06

Description Implementation of a Bayesian approach for estimating a mixture of gamma distributions in which the mixing occurs over the shape parameter. This family provides a flexible and novel approach for modeling heavy-tailed distributions, it is computationally efficient, and it only requires to specify a prior distribution for a single parameter.

Author Sergio Venturini

Maintainer Sergio Venturini <sergio.venturini@unibocconi.it>

Depends gtools, methods, utils

Imports graphics, grDevices, stats

License GPL (>= 2)

URL <http://projecteuclid.org/euclid.aoas/1215118537>

LazyLoad yes

NeedsCompilation no

Repository CRAN

Date/Publication 2015-07-08 00:27:06

R topics documented:

GSM-package	2
allcurves.q	2
estim.gsm	3
estim.gsm_theta	4
gsm-class	6
GSMDist	7
plot-methods	8
predict-methods	9
summary-methods	10

Index	12
--------------	-----------

GSM-package

Estimation of a Gamma Shape Mixture Model

Description

This package implements a Bayesian approach for estimation of a mixture of gamma distributions in which the mixing occurs over the shape parameter. This family provides a flexible and novel approach for modeling heavy-tailed distributions, it is computationally efficient, and it only requires to specify a prior distribution for a single parameter. See Venturini et al. (2008).

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm](#), [estim.gsm.theta](#).

allcurves.q

Utility Function

Description

Utility function for plotting a Gamma Shape Mixture Model density.

Usage

```
allcurves.q(post, perc)
```

Arguments

post	matrix containing of a mixture's density posterior draws.
perc	percentile, a value that satisfies $0 < \text{perc} < 1$.

Details

This is a utility function used to generate the credibility bands for a Gamma Shape Mixture density within [plot](#).

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

See Also

[plot-methods.](#)

estim.gsm

Estimation of a Gamma Shape Mixture Model (GSM) with collapsing

Description

This function provides the inferential algorithm to estimate a mixture of gamma distributions in which the mixing occurs over the shape parameter. It implements the collapsing approach for the GSM model, as discussed in Venturini et al. (2008).

Usage

```
estim.gsm(y, J, G = 100, M = 600, a, b, alpha, init = list(rep(1 / J, J), NA,
  rep(1, N)))
```

Arguments

y	vector of data.
J	number of mixture components.
G	number of points where to evaluate the GSM density.
M	number of MCMC runs.
a	hyperparameter of the rate parameter prior distribution.
b	hyperparameter of the rate parameter prior distribution.
alpha	hyperparameter of the mixture's weights prior distribution.
init	initialization values.

Details

Suggestions on how to choose J, a and b are provided in Venturini et al. (2008). In that work the alpha vector is always set at $(1/J, \dots, 1/J)$, but here one is free to choose the value of the generic element of alpha.

Value

estim.gsm returns an object of `class` "gsm", which is a list with the following components:

fdens	matrix containing the posterior draws for the mixture's density.
theta	vector containing the posterior draws for the mixture's rate parameter.
weight	matrix containing the posterior draws for the mixture's weights.
label	matrix containing the posterior draws for the mixture's labels.
data	vector of data.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm_theta](#), [summary-methods](#), [plot-methods](#).

Examples

```
## Not run:
set.seed(2040)
y <- rgsm(500, c(.1, .3, .4, .2), 1)
burnin <- 100
mcmcsim <- 500
J <- 250
gsm.out <- estim.gsm(y, J, 300, burnin + mcmcsim, 6500, 340, 1/J)
summary(gsm.out, plot = TRUE, start = (burnin + 1))
plot(gsm.out, ndens = 0, nbin = 20, histogram = TRUE, start = (burnin + 1))
## End(Not run)
```

estim.gsm_theta

Estimation of a Gamma Shape Mixture Model (GSM)

Description

This function provides the inferential algorithm to estimate a mixture of gamma distributions in which the mixing occurs over the shape parameter. It implements the standard approach for the GSM model, as discussed in Venturini et al. (2008).

Usage

```
estim.gsm_theta(y, J, G = 100, M = 600, a, b, alpha, init = list(rep(1 / J, J),
  J / max(y), rep(1, N)))
```

Arguments

y	vector of data.
J	number of mixture components.
G	number of points where to evaluate the GSM density.
M	number of MCMC runs.
a	hyperparameter of the rate parameter prior distribution.

b	hyperparameter of the rate parameter prior distribution.
alpha	hyperparameter of the mixture's weights prior distribution.
init	initialization values.

Details

Suggestions on how to choose J, a and b are provided in Venturini et al. (2008). In that work the alpha vector is always set at $(1/J, \dots, 1/J)$, but here one is free to choose the value of the generic element of alpha.

Value

estim.gsm_theta returns an object of class "gsm", which is a list with the following components:

fdens	matrix containing the posterior draws for the mixture's density.
theta	vector containing the posterior draws for the mixture's rate parameter.
weight	matrix containing the posterior draws for the mixture's weights.
label	matrix containing the posterior draws for the mixture's labels.
data	vector of data.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm](#), [summary-methods](#), [plot-methods](#).

Examples

```
## Not run:
set.seed(2040)
y <- rgsm(500, c(.1, .3, .4, .2), 1)
burnin <- 100
mcmcsim <- 500
J <- 250
gsm.out <- estim.gsm_theta(y, J, 300, burnin + mcmcsim, 6500, 340, 1/J)
summary(gsm.out, plot = TRUE, start = (burnin + 1))
plot(gsm.out, ndens = 0, nbin = 20, histogram = TRUE, start = (burnin + 1))
## End(Not run)
```

gsm-class

Class "gsm". Result of Gamma Shape Mixture Estimation.

Description

This class encapsulates results of a Gamma Shape Mixture estimation procedure.

Objects from the Class

Objects can be created by calls of the form `new("gsm", fdens, theta, weight, data)`, but most often as the result of a call to `estim.gsm` or `estim.gsm_theta`.

Slots

`fdens`: Object of class "matrix"; posterior draws from the MCMC simulation algorithm of the Gamma Shape Mixture density.

`theta`: Object of class "numeric"; posterior draws from the MCMC simulation algorithm of the Gamma Shape Mixture scale parameter.

`weight`: Object of class "matrix"; posterior draws from the MCMC simulation algorithm of the Gamma Shape Mixture weights.

`label`: Object of class "matrix"; posterior draws from the MCMC simulation algorithm of the Gamma Shape Mixture labels.

`data`: Object of class "numeric"; original data.

Methods

plot signature(x = "gsm", y = "missing"): Plot Gamma Shape Mixture estimate.

predict signature(object = "gsm"): Estimate of the Gamma Shape Mixture upper tail.

summary signature(object = "gsm"): Generate object summary.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm](#), [summary-methods](#), [plot-methods](#), [predict-methods](#), [summary-methods](#).

 GSMDist

Utility Function

Description

Function evaluations for a Gamma Shape Mixture Model.

Usage

```

dgsm(x, weight, rateparam)
pgsm(q, weight, rateparam, lower.t = TRUE)
rgsm(n, weight, rateparam)
qgsm(p, x = NULL, weight, rateparam, alpha = .05, br = c(0, 1000), lower.t = TRUE)

```

Arguments

x, q	vector of quantiles.
n	number of observations.
p	vector of probabilities.
weight	vector of mixture weights.
rateparam	reciprocal of the shape parameter, as in GammaDist .
alpha	outside the interval (alpha, 1 - alpha) the quantiles are found by searching for the root of $F(x) - p = 0$.
br	a vector containing the end-points of the interval to be searched for the root.
lower.t	logical; if TRUE (default), probabilities are $P[X \leq x]$ otherwise, $P[X > x]$.

Details

The parametrisation implemented in this function is described in Venturini et al. (2008).

Value

[dgsm](#) gives the density, [pgsm](#) gives the distribution function, [qgsm](#) gives the quantile function, and [rgsm](#) generates random deviates.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[dgamma](#), [pgamma](#), [rgamma](#), [uniroot](#).

plot-methods

Plot of a Gamma Shape Mixture Model

Description

plot method for class "gsm". This function plots the output of a Gamma Shape Mixture estimation procedure.

Usage

```
## S4 method for signature 'gsm,missing'
plot(x, ndens = 5, xlab = "x", ylab = "density", nbin = 10,
     histogram = FALSE, bands = FALSE, confid = .95, start = 1, ...)
```

Arguments

x	object of class "gsm"; a list returned by the estim.gsm or estim.gsm_theta functions.
ndens	number of simulated density curves to plot.
xlab	a title for the x axis.
ylab	a title for the y axis.
nbin	number of bins for the histogram.
histogram	logical; if TRUE the histogram is plotted on the figure.
bands	logical; if TRUE the 95% credibility bands are overlaid on the density graph.
confid	confidence level for the pointwise credibility bands around the density estimate.
start	MCMC run to start from.
...	further arguments passed to or from other methods.

Details

To produce a standard histogram with the estimated density curve superimposed on it, simply set `ndens` to 0 and `histogram` to TRUE.

Value

List with the following components:

xval	horizontal coordinates.
yval	vertical coordinates (pointwise density posterior means).

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm](#), [estim.gsm_theta](#), [summary-methods](#), [predict-methods](#).

Examples

```
set.seed(2040)
y <- rgsm(500, c(.1, .3, .4, .2), 1)
burnin <- 5
mcmcsim <- 10
J <- 250
gsm.out <- estim.gsm(y, J, 300, burnin + mcmcsim, 6500, 340, 1/J)
par(mfrow = c(3, 2))
plot(gsm.out)
plot(gsm.out, ndens = 0, nbin = 20, start = (burnin + 1))
plot(gsm.out, ndens = 0, nbin = 20, histogram = TRUE, start = (burnin + 1))
plot(gsm.out, ndens = 0, nbin = 20, histogram = TRUE, bands = TRUE, start = (burnin + 1))
plot(gsm.out, ndens = 5, nbin = 20, histogram = TRUE, bands = TRUE, start = (burnin + 1))
plot(gsm.out, ndens = 0, nbin = 20, bands = TRUE, start = (burnin + 1))
```

predict-methods

Tail Probability Estimation for a Gamma Shape Mixture Model

Description

predict method for class "gsm". This function allows to estimate the tail probability of a Gamma Shape Mixture Model using the output of the [estim.gsm](#) or [estim.gsm_theta](#) procedures.

Usage

```
## S4 method for signature 'gsm'
predict(object, thresh, start = 1, ...)
```

Arguments

object	object of class "gsm"; a list returned by the estim.gsm or estim.gsm_theta functions.
thresh	threshold value.
start	MCMC run to start from.
...	further arguments passed to or from other methods.

Details

The tail probability is estimated by applying the standard Rao-Blackwellized estimator on the Gibbs sampling realizations obtained through the `estim.gsm` or `estim.gsm_theta` procedures.

Value

A numerical vector containing the posterior draws for the tail probability exceeding the value of `thresh`.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

`estim.gsm`, `estim.gsm_theta`, `predict-methods`, `plot-methods`.

Examples

```
set.seed(2040)
y <- rgsm(500, c(.1, .3, .4, .2), 1)
burnin <- 5
mcmcsim <- 10
J <- 250
gsm.out <- estim.gsm(y, J, 300, burnin + mcmcsim, 6500, 340, 1/J)
thresh <- c(0.1, 0.5, 0.75, 1, 2)
tail.prob.est <- tail.prob.true <- rep(NA, length(thresh))
for (i in 1:length(thresh)){
  tail.prob.est[i] <- mean(predict(gsm.out, thresh[i]))
  tail.prob.true[i] <- sum(y > thresh[i])/length(y)
}
qqplot(tail.prob.true, tail.prob.est, main = "Q-Q plot of true vs. estimated tail probability")
abline(0, 1, lty = 2)
```

Description

summary method for class "gsm". This function allows to summarize the output of a Gamma Shape Mixture estimate procedure like `estim.gsm` or `estim.gsm_theta`.

Usage

```
## S4 method for signature 'gsm'  
summary(object, plot = FALSE, start = 1, ...)
```

Arguments

object	object of class "gsm"; a list returned by the <code>estim.gsm</code> or <code>estim.gsm_theta</code> functions.
plot	logical; if TRUE produces a bar plot of the mixture weights posterior means.
start	MCMC run to start from.
...	further arguments passed to or from other methods.

Value

The function `summary` computes and returns a list of summary statistics of the fitted gamma shape mixture given in `object`, in particular

theta	summary index of the theta parameter posterior draws.
weight posterior means	vector of the mixture weights posterior means.

Author(s)

Sergio Venturini <sergio.venturini@unibocconi.it>

References

Venturini, S., Dominici, F. and Parmigiani, G. (2008), "Gamma shape mixtures for heavy-tailed distributions". *Annals of Applied Statistics*, **Volume 2**, Number 2, 756–776. <http://projecteuclid.org/euclid.aoas/1215118537>

See Also

[estim.gsm](#), [estim.gsm_theta](#), [plot-methods](#), [predict-methods](#).

Examples

```
set.seed(2040)  
y <- rgsm(500, c(.1, .3, .4, .2), 1)  
burnin <- 5  
mcmcsim <- 10  
J <- 250  
gsm.out <- estim.gsm(y, J, 300, burnin + mcmcsim, 6500, 340, 1/J)  
summary(gsm.out, TRUE, start = (burnin + 1))
```

Index

- *Topic **class**
 - [gsm-class](#), [6](#)
- *Topic **distribution**
 - [estim.gsm](#), [3](#)
 - [estim.gsm_theta](#), [4](#)
 - [GSM-package](#), [2](#)
 - [GSMDist](#), [7](#)
- *Topic **methods**
 - [gsm-class](#), [6](#)
 - [plot-methods](#), [8](#)
 - [predict-methods](#), [9](#)
 - [summary-methods](#), [10](#)
- *Topic **mixture**
 - [estim.gsm](#), [3](#)
 - [estim.gsm_theta](#), [4](#)
 - [GSM-package](#), [2](#)
- [allcurves.q](#), [2](#)
- [class](#), [3](#), [5](#)
- [dgamma](#), [8](#)
- [dgsms](#), [7](#)
- [dgsms \(GSMDist\)](#), [7](#)
- [estim.gsm](#), [2](#), [3](#), [5](#), [6](#), [8–11](#)
- [estim.gsm_theta](#), [2](#), [4](#), [4](#), [6](#), [8–11](#)
- [GammaDist](#), [7](#)
- [gsm-class](#), [6](#)
- [GSM-package](#), [2](#)
- [GSMDist](#), [7](#)
- [initialize](#), [gsm-method \(gsm-class\)](#), [6](#)
- [pgamma](#), [8](#)
- [pgsm](#), [7](#)
- [pgsm \(GSMDist\)](#), [7](#)
- [plot](#), [2](#)
- [plot](#), [ANY](#), [ANY-method \(plot-methods\)](#), [8](#)
- [plot](#), [gsm](#), [missing-method \(plot-methods\)](#), [8](#)
- [plot-methods](#), [8](#)
- [predict](#), [ANY-method \(predict-methods\)](#), [9](#)
- [predict](#), [gsm-method \(predict-methods\)](#), [9](#)
- [predict-methods](#), [9](#)
- [qgsms](#), [7](#)
- [qgsms \(GSMDist\)](#), [7](#)
- [rgamma](#), [8](#)
- [rgsm](#), [7](#)
- [rgsm \(GSMDist\)](#), [7](#)
- [summary](#), [ANY-method \(summary-methods\)](#), [10](#)
- [summary](#), [gsm-method \(summary-methods\)](#), [10](#)
- [summary-methods](#), [10](#)
- [uniroot](#), [8](#)