Package ‘ocomposition’

August 10, 2015

Type  Package
Title  Regression for Rank-Indexed Compositional Data
Version 1.1
Date  2015-08-08
Author  Arturas Rozenas, New York University
Maintainer  Arturas Rozenas <ar199@nyu.edu>
Description  Regression model where the response variable is a rank-indexed compositional vector (non-negative values that sum up to one and are ordered from the largest to the smallest). Parameters are estimated in the Bayesian framework using MCMC methods.
License  GPL (>= 2)
Depends  R (>= 3.2.1), coda, bayesm
LazyLoad  yes
NeedsCompilation  no
Repository  CRAN
Date/Publication  2015-08-10 09:09:58

R topics documented:

data ................................................................. 2
fitcomp ......................................................... 2
negbinom ....................................................... 3
plot.comphat ................................................... 4
predict.composition .......................................... 4
summary.composition ................................. 5

Index  6
Example data

Description

CSES data (modules I-III). The response variable is a matrix of party vote-shares across democracies. The covariate is the average number of electoral districts in a country.

Usage

data(data)

Format

A list of vote-shares and the list of average electoral districts.

Source

www.cses.org

Gibbs sampler for parameter estimation

Description

The main regression function for compositional rank-index data. For units \( i = 1, \ldots, n \), the response variable is vector \((y_{i1}, \ldots, y_{in})\), where \( \sum_j y_{ij} = 1 \) and \( y_{i1} \geq y_{i2} \geq \ldots \geq y_{in} \) for all \( i \) and \( y_{ij} \in [0, 1] \) for all \( i \) and \( j \). The regression model has two parts: a truncated negative binomial model for the count of non-zero components and a set of seemingly unrelated t regressions for the compositions. See References for further details.

Usage

fitcomp(data.v, data.x, n.formula, v.formula, l.bound = 1, n.sample = 100, burn = 0, thin = 1, init = NULL)

Arguments

data.v Matrix of compositional data: rows for units and columns for components. Rows must add up to 1; if not, they are automatically rescaled. NA values turned into 0 automatically. Ordering done automatically.
data.x Data frame with covariates, missing values not allowed.
n.formula formula for the number of components: e.g., ~ x1 + x2 + factor(z).
v.formula formula for the size of components: e.g., ~ x1 + x2.
### negbinom

1. **bound**
   - lower bound for the negative binomial regression; must be greater or equal to 1; default = 1.

2. **n.sample**
   - number of samples you want to have after burn-in and thinning; default 100

3. **burn**
   - number of burn-in samples; default 0

4. **thin**
   - thinning of the MCMC chain; default 1

5. **init**
   - initial parameters; not required

### Value

- **g**
  - samples of gamma coefficients for the multivariate regression model

- **b**
  - posterior samples of the coefficients for the negative binomial regression

- **mu**
  - hyperparameters for gamma coefficients

- **rho**
  - shrinkage hyperparameters for gamma coefficients

- **Sigma**
  - posterior samples of the covariance matrix

- **nu**
  - degrees of freedom for the Student’s t distribution

### References


### Examples

```r
data(data)
out <- fitcomp(data$v, data$m, ~ log(m) - log(m) + log(n), n.sample = 50)
summary(out)

# predict distribution of votes in a country with 5-member median district
v.hat <- predict(out, data.frame(m=5))
plot(v.hat)
```

---

Truncated negative binomial distribution function and sampling.

### Description

Distribution function of truncated negative binomial distribution and random draws from the distribution (uses rejection sampling).

### Usage

```r
dtnegbin(x, mu, dispersion, l.bound)
rtnegbin(N, mu, dispersion, l.bound)
```
Arguments

\begin{itemize}
  \item \texttt{x} \hspace{1cm} \text{value at which density is evaluated}
  \item \texttt{mu} \hspace{1cm} \text{mean of the distribution}
  \item \texttt{dispersion} \hspace{1cm} \text{dispersion parameter}
  \item \texttt{lbound} \hspace{1cm} \text{the lower bound of truncation}
  \item \texttt{N} \hspace{1cm} \text{Number of draws}
\end{itemize}

\texttt{plot.comphat} \hspace{1cm} \textit{Plot predicted composition.}

Description

Plots predicted components with 95\% confidence intervals.

Usage

\begin{verbatim}
## S3 method for class 'comphat'
plot(x, ...)
\end{verbatim}

Arguments

\begin{itemize}
  \item \texttt{x} \hspace{1cm} \text{Object of class comphat generated by predict.composition.}
  \item \texttt{...}
\end{itemize}

\texttt{predict.composition} \hspace{1cm} \textit{Predicted compositional vector}

Description

Computes the predicted composition given a vector of covariates \texttt{newdata}.

Usage

\begin{verbatim}
## S3 method for class 'composition'
predict(object, newdata, n.method = "median", lbound = NULL, ...)
\end{verbatim}

Arguments

\begin{itemize}
  \item \texttt{object} \hspace{1cm} \text{Object of class composition from the previous call of \texttt{fitcomp}.}
  \item \texttt{newdata} \hspace{1cm} \text{data.frame of predictors.}
  \item \texttt{n.method} \hspace{1cm} \text{If "median" (default), then the number of components is set to the median of the truncated negative binomial distribution, conditional on the estimated parameters; alternatively, it can be set to "mode."}
  \item \texttt{lbound} \hspace{1cm} \text{Set the lower bound for predicted number of components. If NULL (default), then \texttt{lbound} is taken from the \texttt{composition} object.}
  \item \texttt{...} \hspace{1cm} \text{additional arguments}
\end{itemize}
Value

Matrix of predicted probabilities of for each component. Rows are MCMC iterations and columns are the components.

Description

Summarizes the basic properties of the MCMC output: means, standard deviations, MCMC errors, and Heidel convergence diagnostics.

Usage

```r
## S3 method for class 'composition'
summary(object, individual = NULL, ...)
```

Arguments

- `object` MCMC output list member of class `composition`, generated by call to `fitcomp`.
- `individual` Integer from 1 to the number of components in data. Outputs summary for that component; if `NULL`, then outputs summaries of the hyperparameters.
- `...`
Index

*Topic `\textasciitilde kwd1`
  plot.comphat, 4
  predict.composition, 4
  summary.composition, 5

*Topic `\textasciitilde kwd2`
  plot.comphat, 4
  predict.composition, 4
  summary.composition, 5

*Topic `compositional`
  fitcomp, 2

*Topic `datasets`
  data, 2

*Topic `multivariate response`
  fitcomp, 2

  data, 2
  dtnegbin (negbinom), 3

  fitcomp, 2
  negbinom, 3

  plot.comphat, 4
  predict.composition, 4

  rtnegbin (negbinom), 3
  summary.composition, 5