Package: mbsts (via r-universe)

August 21, 2024

Type Package

Title Multivariate Bayesian Structural Time Series

Version 3.0

Author Jinwen Qiu <qjwsnow_ctw@hotmail.com>, Ning Ning <patricianing@gmail.com>

Maintainer Ning Ning <patricianing@gmail.com>

Description Tools for data analysis with multivariate Bayesian structural time series (MBSTS) models. Specifically, the package provides facilities for implementing general structural time series models, flexibly adding on different time series components (trend, season, cycle, and regression), simulating them, fitting them to multivariate correlated time series data, conducting feature selection on the regression component.

License LGPL-2.1

Encoding UTF-8

NeedsCompilation no

Depends R(>= 3.5.0)

Imports KFAS, Matrix, pscl, MCMCpack, MASS, ggplot2, reshape2, BBmisc, matrixStats, methods

RoxygenNote 7.2.1

Collate 'mbsts.forecast.R' 'mbsts_class.R' 'mbsts_function.R' 'package.R' 'para.est.R' 'plot_comp.R' 'plot_cvg.R' 'plot_prob.R' 'sim_data.R' 'tsc.setting.R'

Date/Publication 2023-01-07 01:10:02 UTC

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

Config/testthat/edition 3

VignetteBuilder knitr

Repository https://patning.r-universe.dev

RemoteUrl https://github.com/cran/mbsts

RemoteRef HEAD

RemoteSha 141b239678dbf3d2afa92894ec3037859b88faf9

Contents

`mbsts-package` *Multivariate Bayesian Structural Time Series*

Description

Tools for data analysis with multivariate Bayesian structural time series (MBSTS) models. Specifically, the package provides facilities for implementing general structural time series models, flexibly adding on different time series components (trend, season, cycle, and regression), simulating them, fitting them to multivariate correlated time series data, conducting feature selection on the regression component.

Documentation

mbsts is described in Ning and Qiu (2021).

License

mbsts is provided under the LGPL-2.1 License.

Author(s)

Jinwen Qiu, Ning Ning

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

`mbsts-class` *Constructor for the MBSTS model*

Description

This class constructor build an object of MBSTS class, encoding a state space model together with a uni- or multi-variate time series, which is central to all the package's functionality. One implements the MBSTS model by specifying some or all of its basic components.

Slots

`Xtrain` A $(n * K)$-dimensional matrix containing all candidate predictor series for each target series. $K = \sum k_i$ is the number of all candidate predictors for all target series. The first k_1 variables are the set of candidate predictors for the first target series, and the next k_2 variables are the set of candidate predictors for the second target series, etc. Note that, one variable can appear in the X.star several times, since different target series can contain the same candidate predictors.

`Ind` A ($K*$(mc-burn))-dimensional matrix containing MCMC draws of the indicator variable. If X.star is null, it will not be returned.

`beta.hat` A ($K*$(mc-burn))-dimensional matrix containing MCMC draws of regression coefficients. If X.star is null, it will not be returned.

`B.hat` A ($K * m*$(mc-burn))-dimensional array generated by combining beta.hat for all target series. If X.star is null, it will not be returned.

`ob.sig2` A ($m*m*$(mc-burn))-dimensional array containing MCMC draws of variance-covariance matrix for residuals.

`States` A ($n * m1*$(mc-burn))-dimensional array containing MCMC draws of all time series components, where $m1$ is the number of all time series components. If the STmodel is null, it will not be returned.

`st.sig2` A ($K*$(mc-burn))-dimensional matrix containing MCMC draws of variances for time series components. If the STmodel is null, it will not be returned.

`ki` A vector of integer values denoting the acumulated number of predictors for target series. For example, if there are three target series where the first has 8 predictors, the second has 6 predictors, and the third has 10 predictors, then the vector is c$(8, 14, 24)$.

`ntrain` A numerical value for number of observations.

`mtrain` A numerical value for number of response variables.

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

`mbsts.forecast` *Specification of time series components*

Description

Generate draws from the posterior predictive distribution of a mbsts object. Samples from the posterior predictive distribution of the MBSTS model.

Usage

```
mbsts.forecast(object, STmodel, newdata, steps = 1)
```

Arguments

`object`	An object of the mbsts class created by a call to the mbsts_function function.
`STmodel`	An object of the SSModel class created by a call to the tsc.setting function.
`newdata`	A vector or matrix containing the predictor variables to use in making the prediction. This is only required if the mbsts model has a regression component.
`steps`	An integer value describing the number of time steps ahead to be forecasted. If it is greater than the number of new observations in the newdata, zero values will fill in missing new observations.

Value

An object of predicted values which is a list containing the following:

`pred.dist`	An array of draws from the posterior predictive distribution. The first dimension in the array represents time, the second dimension denotes each target series, and the third dimension indicates each MCMC draw.
`pred.mean`	A matrix giving the posterior mean of the prediction for each target series.
`pred.sd`	A matrix giving the posterior standard deviation of the prediction for each target series.
`pred.se`	A matrix giving the posterior standard error of the prediction for each target series, calculated by pred.sd divided by the square root of the numer of MCMC iterations.

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

`mbsts_function` *Main function for the multivariate Bayesian structural time series (MBSTS) model*

Description

The MBSTS model uses MCMC to sample from the posterior distribution of a MBSTS model. The model is given by

$$y = \mu + \tau + \omega + \beta X + \epsilon,$$

where μ, τ, ω, βX, and ϵ denote the trend component, the seasonal component, the cycle component, the regression component, and the error term, respectively. Note that, without a regression component, the MBSTS model is an ordinary state space time series model. The predictors and response variables in the MBSTS model are designed to be contemporaneous. Lags and differences can be generated by manipulating the predictor matrix. The "spike-and-slab" prior is used for the regression component of models, which enables feature selection among a large number of features.

Usage

```
mbsts_function(
  Y,
  Xtrain,
  STmodel,
  ki,
  pii,
  b = NULL,
  v0,
  kapp = 0.01,
  R2 = 0.8,
  v = 0.01,
  ss = 0.01,
  mc = 500,
  burn = 50
)

## S4 method for signature 'array'
mbsts_function(
```

```
  Y,
  Xtrain,
  STmodel,
  ki,
  pii,
  b = NULL,
  v0,
  kapp = 0.01,
  R2 = 0.8,
  v = 0.01,
  ss = 0.01,
  mc = 500,
  burn = 50
)
```

Arguments

Y	A $(n * m)$-dimensional matrix containing multiple target series, where n is the number of observations and m is the number of target series.
Xtrain	A $(n * K)$-dimensional matrix containing all candidate predictor series for each target series. $K = \sum k_i$ is the number of all candidate predictors for all target series. The first k_1 variables are the set of candidate predictors for the first target series, and the next k_2 variables are the set of candidate predictors for the second target series, etc. Note that, one variable can appear in the X.star several times, since different target series can contain the same candidate predictors.
STmodel	A state space model of SSmodel class returned by tsc.setting.
ki	A vector of integer values denoting the acumulated number of predictors for target series. For example, if there are three target series where the first has 8 predictors, the second has 6 predictors, and the third has 10 predictors, then the vector is c$(8, 14, 24)$.
pii	A vector describing the prior inclusion probability of each candidate predictor.
b	NULL or a vector describing the prior means of regression coefficients. The default value is NULL.
v0	A numerical value describing the prior degree of freedom of the inverse Wishart distribution for Σ_ϵ.
kapp	A scalar value describing the number of observations worth of weight on the prior mean vector. The default value is 0.01.
R2	A numerical value taking value in $[0, 1]$, describing the expected percentage of variation of Y to be explained by the model. The default value is 0.8.
v	A numerical value describing the prior degree of freedom of the inverse Wishart distribution for $(\Sigma_\mu, \Sigma_\delta, \Sigma_\tau, \Sigma_\omega)$. The default value is 0.01.
ss	A numerical value describing the prior scale matrix of the inverse Wishart distribution for $(\Sigma_\mu, \Sigma_\delta, \Sigma_\tau, \Sigma_\omega)$. The default value is 0.01.
mc	A positive integer giving the desired number of MCMC draws. The default value is 500.
burn	A positive integer giving the number of initial MCMC draws to be discarded. The default value is 50.

Value

An object of mbsts class

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

---

para.est *Regression parameter estimation by the MBSTS Model*

Description

Generate feature selection and parameter estimation results of a mbsts object. Provide means and standard deviations of parameter estimation results for selected features.

Usage

```
para.est(object, prob.threshold = 0.2)

## S4 method for signature 'mbsts'
para.est(object, prob.threshold = 0.2)
```

Arguments

object	An object of the mbsts class created by a call to the mbsts_function function.
prob.threshold	A numerical value used as the threshold to only include predictors whose inclusion probabilities are higher than it in the plot. The default is 0.2.#' @param prob.threshold A numerical value used as the threshold to only include predictors whose inclusion probabilities are higher than it in the plot. The default value is 0.2.

Value

A list with the following components

index	An array of feature selection results.
para.est.mean	An array of means of parameter estimation values of selected features.
para.est.sd	An array of standard deviations of parameter estimation values of selected features.

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

---

plot_comp *Plot Posterior State Components*

Description

Plots of the mean of posterior state components of each target series, which is generated by the model training procedure of the MBSTS model.

Usage

```
plot_comp(
  object,
  slope,
  local,
  season,
  cyc,
  time = NULL,
  title = NULL,
  component_selection = "All"
)

## S4 method for signature 'mbsts'
plot_comp(
  object,
```

```
  slope,
  local,
  season,
  cyc,
  time = NULL,
  title = NULL,
  component_selection = "All"
)
```

Arguments

object	An object of the mbsts class created by a call to the mbsts_function function.
slope	A logical vector indicating whether there is trend for each target series, such as c(T,T).
local	A logical vector indicating whether there is local level for each target series, such as c(T,T).
season	A numerical vector indicating the seasonality for each target series, such as c(12,0).
cyc	A logical vector indicating whether there is a cycle component for each target series, such as c(F,T).
time	Null or a data frame for time index of the time series. The default value is data.frame(seq(1,n)).
title	NULL or a character vector whose entries are titles for the plots of target series' posterior state components, such as c("Posterior State Components of y1", "Posterior State Components of y2"). The default is c("y1","y2",...).
component_selection	A character variable whose value must be one of "All", "Trend", "Seasonal", "Cycle", and "Regression". Here, "Trend" means the trend component only and "All" means all the components.

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

plot_cvg *Plot for Convergence Diagnosis*

Description

Plot of the parameter draw for MCMC iterations after burn-in.

Usage

```
plot_cvg(
  object,
  index,
  type = "o",
  col = "blue",
  pch = 16,
  lty = 1,
  xlab = "Number of iterations",
  ylab = "Estimation",
  main = "Predictor",
  cex.axis = 1.15
)

## S4 method for signature 'mbsts'
plot_cvg(
  object,
  index,
  type = "o",
  col = "blue",
  pch = 16,
  lty = 1,
  xlab = "Number of iterations",
  ylab = "Estimation",
  main = "Predictor",
  cex.axis = 1.15
)
```

Arguments

object	An object of the mbsts class created by a call to the mbsts_function function.
index	A numerical value indicating which predictor to analyze. The index can be generated by a call to the para.est function
type	NULL or a character vector whose entries are titles for the plots of target series' posterior state components, such as c("Posterior State Components of y1", "Posterior State Components of y2"). The default is c("y1","y2",...).
col	The same setting as that of the plot function in the base package.
pch	The same setting as that of the plot function in the base package.

lty	The same setting as that of the plot function in the base package.
xlab	The same setting as that of the plot function in the base package.
ylab	The same setting as that of the plot function in the base package.
main	The same setting as that of the plot function in the base package.
cex.axis	The same setting as that of the plot function in the base package.

Author(s)

Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

---

plot_prob *Plot Inclusion Probabilities*

Description

Plots of the empirical inclusion probabilities for predictors of each target series, based on a user-defined threshold probability. For example, one predictor is selected 100 times in 200 MCMC draws (after discard burn-in draws), the empirical inclusion probability for that predictor is 0.5. If the user-defined threshold probability less than or equal to 0.5, then this predictor will show in the plot.

Usage

```
plot_prob(object, title = NULL, prob.threshold = 0.2, varnames = NULL)

## S4 method for signature 'mbsts'
plot_prob(object, title = NULL, prob.threshold = 0.2, varnames = NULL)
```

Arguments

object	An object of the mbsts class created by a call to the mbsts_function function.
title	NULL or A character vector whose entries are titles for the inclusion probability plots generated for each target series, such as c("Inclusion Probabilities for y1", "Inclusion Probabilities for y2"). If Null, the output is c("y1","y2",...).

prob.threshold	A numerical value used as the threshold to only include predictors whose inclusion probabilities are higher than it in the plot. The default value is 0.2.
varnames	NULL or A character vector whose entries are the variable names for predictors, such as c("x1", "x2", "x3", "x4", "x5", "x6", "x7", "x8", "x1", "x2", "x3", "x4", "x5", "x6", "x7", "x8"). If Null, the output is c("x11","x12",...,"x21","x22",...).

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

---

sim_data *Simulate data*

Description

Generate simulated data in the form of structural time series

Usage

```
sim_data(
  X,
  beta,
  cov,
  k,
  mu,
  rho,
  mean_trend = 1,
  sd_trend = 0.5,
  mean_season = 20,
  sd_season = 0.5,
  mean_cycle = 20,
  sd_cycle = 0.5,
  Dtilde,
  Season,
  vrho,
  lambda
```

```
)

## S4 method for signature 'array'
sim_data(
  X,
  beta,
  cov,
  k,
  mu,
  rho,
  mean_trend = 1,
  sd_trend = 0.5,
  mean_season = 20,
  sd_season = 0.5,
  mean_cycle = 20,
  sd_cycle = 0.5,
  Dtilde,
  Season,
  vrho,
  lambda
)
```

Arguments

X	A $(n * K)$-dimensional matrix containing predictors, where n is the number of observations. $K = \sum k_i$ is the number of all candidate predictors for all target series. The first k_1 variables are the set of candidate predictors for the first target series, and the next k_2 variables are the set of candidate predictors for the second target series, etc.
beta	A $(K * m)$-dimensional matrix containing all candidate predictor series for each target series.
cov	A $(m * m)$-dimensional matrix containing covariances
k	A m-dimensional array containing the number of candidate predictors for each of the m target series.
mu	A m-dimensional array with 1 representing modeling with trend for this target time series.
rho	A m-dimensional array representing the learning rates at which the local trend is updated.
mean_trend	A numerical value standing for the mean of the error term of the trend component. The default value is 1.
sd_trend	A numerical value standing for the standard deviation of the error term of the trend component. The default value is 0.5.
mean_season	A numerical value standing for the mean of the error term of the seasonal component. The default value is 20.
sd_season	A numerical value standing for the standard deviation of the error term of the seasonal component. The default value is 0.5.

mean_cycle	A numerical value standing for the mean of the error term of the cycle component. The default value is 20.
sd_cycle	A numerical value standing for the standard deviation of the error term of the cycle component. The default value is 0.5.
Dtilde	A m-dimensional array with 1 representing level in the trend component.
Season	A m-dimensional array indicating the seasonality for each target series, such as c(12,0).
vrho	A m-dimensional array of the decay value parameter of the cycle component for each target series, such as c(0,0.99).
lambda	A m-dimensional array of the frequence parameter of the cycle component for each target series, such as c(0,pi/100).

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

Examples

```
###############Setup###########
n<-505 #n: sample size
m<-2 #m: dimension of target series

cov<-matrix(c(1.1,0.7,0.7,0.9), nrow=2, ncol=2) #covariance matrix of target series

###############Regression component###########
#coefficients for predictors
beta<-t(matrix(c(2,-1.5,0,4,2.5,0,0,2.5,1.5,-1,-2,0,0,-3,3.5,0.5),nrow=2,ncol=8))

set.seed(100)
X1<-rnorm(n,5,5^2)
X4<-rnorm(n,-2,5)
X5<-rnorm(n,-5,5^2)
X8<-rnorm(n,0,100)
X2<-rpois(n, 10)
X6<-rpois(n, 15)
X7<-rpois(n, 20)
X3<-rpois(n, 5)
X<-cbind(X1,X2,X3,X4,X5,X6,X7,X8)
```

```
###############Simulated data################
set.seed(100)
data=sim_data(X=X, beta=beta, cov, k=c(8,8), mu=c(1,1), rho=c(0.6,0.8),
              Dtilde=c(-1,3), Season=c(100,0), vrho=c(0,0.99), lambda=c(0,pi/100))
```

`tsc.setting` *Specification of time series components*

Description

Specify three time series components for the MBSTS model: the generalized linear trend component, the seasonal component, and the cycle component.

Usage

```
tsc.setting(Ytrain, mu, rho, S, vrho, lambda)
```

Arguments

`Ytrain`	The multivariate time series to be modeled.
`mu`	A vector of logic values indicating whether to include a local trend for each target series.
`rho`	A vector of numerical values taking values in $[0, 1]$, describing the learning rates at which the local trend is updated for each target series. The value 0 in the j-th entry indicates that the j-th target series does not include slope of trend.
`S`	A vector of integer values representing the number of seasons to be modeled for each target series. The value 0 in the j-th entry indicates that the j-th target series does not include the seasonal component.
`vrho`	A vector of numerical values taking values in $[0, 1]$, describing a damping factor for each target series. The value 0 in the j-th entry indicates that the j-th target series does not include the cycle component.
`lambda`	A vector of numerical values, whose entries equal to $2\pi/q$ with q being a period such that $0 < \lambda < \pi$, describing the frequency.

Value

An object of the SSModel class.

Author(s)

Jinwen Qiu <qjwsnow_ctw@hotmail.com> Ning Ning <patricianing@gmail.com>

References

Qiu, Jammalamadaka and Ning (2018), Multivariate Bayesian Structural Time Series Model, *Journal of Machine Learning Research* 19.68: 1-33.

Ning and Qiu (2021), The mbsts package: Multivariate Bayesian Structural Time Series Models in R.

Jammalamadaka, Qiu and Ning (2019), Predicting a Stock Portfolio with the Multivariate Bayesian Structural Time Series Model: Do News or Emotions Matter?, *International Journal of Artificial Intelligence*, Vol. 17, Number 2.

Index