

Package ‘mbsts’

January 9, 2020

Type Package

Title Multivariate Bayesian Structural Time Series

Version 1.0

Author Jinwen Qiu <qjwsnow_ctw@hotmail.com>

Maintainer Jinwen Qiu <qjwsnow_ctw@hotmail.com>

Description

Multivariate time series regression using dynamic linear models fit by MCMC. See Qiu, Jammalamadaka and Ning (2018) <<http://www.jmlr.org/papers/volume19/18-009/18-009.pdf>>.

License LGPL-2.1

NeedsCompilation no

Depends pscl, MASS, KFAS, MCMCpack, Matrix, R(>= 3.5)

Repository CRAN

Date 2020-01-07

Date/Publication 2020-01-09 18:10:05 UTC

R topics documented:

exdata	1
mbsts	2
mbsts.forecast	5
tsc.setting	7

Index	10
--------------	-----------

exdata	<i>Example Data</i>
--------	---------------------

Description

Data generated by dynamic linear model including 2 target series (one with linear trend and seasonality; another with generalized trend and cycle) and 16 contemporaneous predictors.

Usage

```
data(exdata)
```

Format

A data frame with 500 observations on the following 18 variables. First two variables are 2 target series. The next 8 variables are candidate predictors for 1st target series, while the rest of them are candidate predictors for 2nd target series.

Source

See Equation (40) in the paper wrote by Jamalamadaka, Qiu and Ning (2018) <<https://arxiv.org/pdf/1801.03222.pdf>>

Examples

```
data(exdata)
ts.plot(exdata[, 1:2])
```

mbsts

Multivariate Bayesian Structural Time Series

Description

Uses MCMC to sample from the posterior distribution of a Multivariate Bayesian structural time series model. This function can be used either with or without contemporaneous predictor variables (in a time series regression).

If predictor variables are present, the regression coefficients are fixed. The predictors and response in the formula are contemporaneous, so if you want lags and differences you need to put them in the predictor matrix yourself. If no predictor variables are used, then the model is an ordinary state space time series model.

The model allows for several useful extensions beyond standard Bayesian dynamic linear models. A spike-and-slab prior is used for the (static) regression component of models that include predictor variables. This is especially useful with large numbers of regressor series.

The model in state space form can be written as:

$$y_t = \mu_t + \tau_t + \omega_t + \beta * X_t + rnorm(0, \Sigma_\epsilon)$$

where μ, τ and ω denotes trend, seasonal and cycle component, respectively.

Usage

```
mbsts(Y,
      X.star = NULL,
      STmodel = NULL,
      ki = NULL,
      pii = NULL,
      b = NULL,
```

```

kapp = 0.1,
R2 = 0.8,
v0 = NULL,
v = 0.01,
ss = 0.01,
mc = 500,
burn = 50)

```

Arguments

Y	An n by m matrix contains multiple target series. The n and m is the number of observations and target series, respectively.
X.star	An n by K matrix contains all candidate predictor series for each target series. The $K = \sum k_i$ is the number of all candidate predictors for all target series. The first k_1 variables make up the set of candidate predictors for 1st target series, and the next k_2 variables are candidate predictors for 2nd target series, etc. One variable can appear in the X.star several times, since different target series can contain same candidate predictors.
STmodel	A state space model of SSmodel object returned by tsc.setting.
ki	A vector of integer values, denoting the location of last candidate predictor for each target series, such as $c(k_1, k_1 + k_2, \dots, K)$
pii	A column matrix with length equal to K, describing the prior inclusion probability of each candidate predictor.
b	A column matrix, describing the prior means for regression coefficients.
kapp	A scalar value, describing the number of observations worth of weight on the prior mean vector.
R2	A numerical value from 0 to 1, describing the expected percent of variation of Y to be explained by model.
v0	A numerical value, describing the degree of freedom for prior Inverse Wishart distribution (Σ_ϵ).
v	A numerical value, describing the prior sample size of Inverse Gama distribution ($\Sigma_\mu, \Sigma_\delta, \Sigma_\tau, \Sigma_\omega$).
ss	A numerical value, describing the prior sum of squares of Inverse Gama distribution ($\Sigma_\mu, \Sigma_\delta, \Sigma_\tau, \Sigma_\omega$).
mc	A positive integer giving the desired number of MCMC draws.
burn	A positive integer giving the initial number of MCMC draws to be discarded.

Details

See the paper wrote by Jammalamadaka, Qiu and Ning (2018) <<https://arxiv.org/pdf/1801.03222.pdf>>

Value

An object of class mbsts which is a list with the following components:

Ind	One K by mc-burn matrix, containing MCMC draws of indicator variable (gama). If the X.star is null, it will not be returned.
beta.hat	One K by mc-burn matrix, containing MCMC draws of regression coefficients (beta).If the X.star is null, it will not be returned.
B.hat	This is beta.hat in matrix form. If the X.star is null, it will not be returned.
ob.sig2	An array with dimension (m*m*(mc-burn)),containing MCMC draws of variance and covariance matrix for residuals.
States	An array with dimension (n*m1*(mc-burn)),containing MCMC draws of all time series components. The m1 is the number of all time series components, respectively. If the STmodel is null, it will not be returned.
st.sig2	One K by mc-burn matrix, containing MCMC draws of variance for time series components. If the STmodel is null, it will not be returned.

Author(s)

Jinwen Qiu <jqiu@pstat.ucsb.edu>

References

- Harvey (1990),"Forecasting, structural time series, and the Kalman filter", Cambridge University Press.
- Durbin and Koopman (2001), "Time series analysis by state space methods", Oxford University Press.
- Goerge and McCulloch (1997) "Approaches for Bayesian variable selection", Statistica Sinica pp 339-74.
- Scott and Varian (2014) "Predicting the present with Bayesian structural time series", International Journal of Mathematical Modelling and Numerical Optimisation, 5 (1-2), 4-23.
- Griffiths (2003) "Bayesian inference in the seemingly unrelated regressions model", Computer aided econometrics, pp 263-290, CRC Press.

See Also

[tsc.setting](#)

Examples

```
data(exdata)

#Two target series
Y<-as.matrix(exdata[,1:2])
#Sixteen candidate predictors
X.star<-as.matrix(exdata[,3:18])

#split dataset into training set and test set
n=dim(Y)[1]
ntrain=n-5
Ytrain<-Y[1:ntrain,]
```

```

Xtrain<-X.star[1:ntrain,]
Ytest<-Y[(ntrain+1):n,]
Xtest<-X.star[(ntrain+1):n,]

#Specify time series components
STmodel<-tsc.setting(Ytrain,mu=c(1,1),rho=c(0.6,1),S=c(4,0),
                    vrho=c(0,0.5),lambda=c(0,pi/10))

#prior parameters setting
#gama
ki<- c(8,dim(Xtrain)[2])
pii<- matrix(rep(0.5,dim(Xtrain)[2]),nrow=dim(Xtrain)[2])

#beta
b<-matrix(0,dim(Xtrain)[2])
kapp<-0.01

#v0 and V0 for obs Sigma
R2<-0.8
v0<-5

#State component Sigma
v<-0.01
ss<-0.01

#train a mbsts model
mbsts.model<-mbsts(Ytrain,Xtrain,STmodel,ki,pii,b,kapp,R2,v0,v,ss,mc=15,burn=5)

```

mbsts.forecast	<i>Multi-steps ahead Forecast by Multivariate Bayesian Structural Time Series Model</i>
----------------	---

Description

Generated draws from the posterior predictive distribution of a mbsts object.

Usage

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

Arguments

mbsts	An object of class mbsts created by a call to the function mbsts.
STmodel	An object of class SSMModel created by a call to the function tsc.setting.
newdata	a vector or matrix containing the predictor variables to use in making the prediction. This is only required if object contains 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, the zero values will fill in missing new observations.

Details

Samples from the posterior distribution of a Bayesian structural time series model. This function can be used either with or without contemporaneous predictor variables (in a time series regression).

If predictor variables are present, the regression coefficients are fixed. The predictors and response in the formula are contemporaneous, so if you want lags and differences you need to put them in the predictor matrix yourself.

If no predictor variables are used, then the model is an ordinary state space time series model.

Value

An object of predicted values by mbsts which is a list with the following components:

pred.distribution

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.

Author(s)

Jinwen Qiu <jqiu@pstat.ucsb.edu>

References

Harvey (1990), "Forecasting, structural time series, and the Kalman filter", Cambridge University Press.

Durbin and Koopman (2001), "Time series analysis by state space methods", Oxford University Press.

Goerge and McCulloch (1997) "Approaches for Bayesian variable selection", Statistica Sinica pp 339-74.

Scott and Varian (2014) "Predicting the present with Bayesian structural time series", International Journal of Mathematical Modelling and Numerical Optimisation, 5 (1-2), 4-23.

See Also

[mbsts](#)

Examples

```
data(exdata)

#Two target series
Y<-as.matrix(exdata[,1:2])
#Sixteen candidate predictors
X.star<-as.matrix(exdata[,3:18])

#split dataset into training set and test set
n=dim(Y)[1]
```

```

ntrain=n-5
Ytrain<-Y[1:ntrain,]
Xtrain<-X.star[1:ntrain,]
Ytest<-Y[(ntrain+1):n,]
Xtest<-X.star[(ntrain+1):n,]

#Specify time series components
STmodel<-tsc.setting(Ytrain,mu=c(1,1),rho=c(0.6,1),S=c(4,0),
                    vrho=c(0,0.5),lambda=c(0,pi/10))

#prior parameters setting
#gama
ki<- c(8,dim(Xtrain)[2])
pii<- matrix(rep(0.5,dim(Xtrain)[2]),nrow=dim(Xtrain)[2])

#beta
b<-matrix(0,dim(Xtrain)[2])
kapp<-0.01

#v0 and V0 for obs Sigma
R2<-0.8
v0<-5

#State component Sigma
v<-0.01
ss<-0.01

#train a mbsts model
mbsts.model<-mbsts(Ytrain,Xtrain,STmodel,ki,pii,b,kapp,R2,v0,v,ss,mc=15,burn=5)

#make a 5-steps prediction
output<-mbsts.forecast(mbsts.model,STmodel,newdata=Xtest,steps=3)
#error<-abs(output$pred.mean-Ytest)
#error

```

tsc.setting

Specification of time series components

Description

Specify three time series components for MBSTS model (generalized linear trend, seasonality and cycle). A generalization of the local linear trend model where the slope exhibits stationarity instead of obeying a random walk, is expressed in the form as:

$$\mu_{t+1} = \mu_t + \delta_t + rnorm(0, \Sigma_\mu)$$

$$\delta_{t+1} = D + \rho(\delta_t - D) + rnorm(0, \Sigma_\delta)$$

The seasonal model in the time domain is:

$$\tau_{t+1}^{(i)} = - \sum_{k=0}^{S_i-2} \tau_{t-k}^{(i)} + rnorm(0, \Sigma_\tau)$$

The cycle component is postulated as:

$$\begin{aligned}\omega_{t+1} &= \text{varrho} * \cos(\lambda)\omega_t + \text{varrho} * \sin(\lambda)\omega^*_t + rnorm(0, \Sigma_\omega) \\ \omega^*_{t+1} &= -\text{varrho} * \sin(\lambda)\omega_t + \text{varrho} * \cos(\lambda)\omega^*_t + rnorm(0, \Sigma_\omega)\end{aligned}$$

Usage

```
tsc.setting(Y, mu = NULL, rho = NULL, S = NULL, vrho = NULL, lambda = NULL)
```

Arguments

Y	The multivariate time series to be modeled, as a numeric matrix convertible to xts. This state model assumes Y contain daily data.
mu	A vector of logic values, indicating whether to include a local trend for each target series.
rho	A vector of numerical values, all of which are between 0 and 1, describing the learning rates at which the local trend is updated for each target series. The value 0 in jth entry indicates that jth 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 jth entry indicates that jth target series does not include seasonal component.
vrho	A vector of numerical values between 0 and 1, describing a damping factor for each target series. The value 0 in jth entry indicates that jth target series does not include cycle component.
lambda	A vector of numerical values, each of them is equal to $2\pi/q$, describing the frequency with q being a period such that $0 < \lambda < \pi$.

Details

The function allows users to include different combination of time series components for each target series.

Value

Returns a customized State Space model, which is an object of class SSMModel.

Author(s)

Jinwen Qiu <jqiu@pstat.ucsb.edu>

References

- Harvey (1990), "Forecasting, structural time series, and the Kalman filter", Cambridge University Press.
- Durbin and Koopman (2001), "Time series analysis by state space methods", Oxford University Press.

See Also

[SSModel](#)

Examples

```
data(exdata)
Y<-as.matrix(exdata[,1:2])
STmodel<-tsc.setting(Y,mu=c(1,1),rho=c(0.6,1),S=c(4,0),
                    vrho=c(0,0.5),lambda=c(0,pi/10))
```

Index

*Topic **datasets**

exdata, [1](#)

*Topic **forecast**

mbsts.forecast, [5](#)

*Topic **model**

mbsts, [2](#)

*Topic **state components**

tsc.setting, [7](#)

exdata, [1](#)

mbsts, [2](#), [6](#)

mbsts.forecast, [5](#)

SSModel, [9](#)

tsc.setting, [4](#), [7](#)