Package ‘ForecastCombinations’

November 23, 2015

Type Package
Title Forecast Combinations
Version 1.1
Date 2015-11-22
Author Eran Raviv
Maintainer Eran Raviv <eeraviv@gmail.com>
Description Aim: Supports the most frequently used methods to combine forecasts. Among others: Simple average, Ordinary Least Squares, Least Absolute Deviation, Constrained Least Squares, Variance-based, Best Individual model, Complete subset regressions and Information-theoretic (information criteria based).
Depends quantreg, quadprog, utils
Suggests MASS
License GPL-2
NeedsCompilation no
Repository CRAN
Date/Publication 2015-11-23 12:20:30

R topics documented:

Forecast_comb ................................................. 1
Forecast_comb_all .......................................... 3

Index

Forecast_comb Forecasts combination using regression, robust regression, constrained regression or variance based
**Description**

Combine different forecasts. Use simple average, Ordinary Least Squares (OLS), robust regression, inverse mean squared error (IMSE), constrained least squares (CLS), or simply use the best forecast based on the MSE metric.

**Usage**

```r
Forecast_comb(obs, fhat, fhat_new = NULL, Averaging_scheme = c("simple", "ols", "robust", "cls", "variance based", "best"))
```

**Arguments**

- `obs` Observed series
- `fhat` fhat Matrix of available forecasts. These are used to retrieve the weights. How each forecast should be weighted in the overall combined forecast.
- `fhat_new` Matrix of available forecasts as a test set. Optional, default to NULL.
- `Averaging_scheme` Which averaging scheme should be used?

**Details**

Performs simple forecast averaging where each forecast carries equal weight: \( \frac{1}{p} \) with \( p \) the column dimension of `fhat`. OLS forecast combination is based on

\[
\text{obs}_t = \text{const} + \sum_{i=1}^{p} w_i \text{fhat}_it + e_t,
\]

where `obs` is the observed values and \( \text{fhat}_i \) is the forecast, one out of the \( p \) forecasts available.

Robust regression performs the same, but minimize different loss function, which is less sensitive to outliers (see quantile regression and references therein).

Constrained least squares minimize the sum of squared errors under the restriction that the weights sum up to 1, and that the forecasts themselves are unbiased (no intercept in the regression).

The variance-based method computes the mean squared error and weigh the forecasts according to their accuracy. Accurate forecasts (based on MSE metric) receive relatively more weight.

The best restrict all the weights to zero apart from the best forecast, again based on the MSE. Essentially selecting only one forecast to be used.

**Value**

`Forecast_comb` returns a list that contains the following objects:

- `fitted` Vector of fitted values.
- `pred` Vector of prediction. This object is empty if there was no test matrix `fhat_new` provided.
- `weights` Vector of weights based on the `Averaging_scheme`.
Author(s)

Eran Raviv (<eeraviv@gmail.com>)

References


Examples

library(MASS)

tt <- nrow(boston)/2

TT <- nrow(boston)
y <- boston[, 14] # dependent variable is column 14

# Create two sets of explanatory variables
x1 <- boston[, 1:6] # The first 6 explanatory variables
x2 <- boston[, 7:13] # The last 6 explanatory variables

# Create two forecasts based on the two different x1 and x2
coef1 <- lm(y ~ as.matrix(x1))$coef
coef2 <- lm(y ~ as.matrix(x2))$coef
f1 <- t(coef1 %*% t(cbind(rep(1, tt), boston[(tt+1):TT, 1:6])))
f2 <- t(coef2 %*% t(cbind(rep(1, tt), boston[(tt+1):TT, 7:13])))
ff <- cbind(f1, f2)
scheme <- c("simple", "ols", "robust", "variance based", "cls", "best")

example0 <- list()

for (i in scheme) {
  example0[[i]] <- Forecast_comb(obs = boston[(tt+1):TT, 14], fhat = ff, Averaging_scheme = i)
  cat(i, ": ", sqrt(mean((example0[[i]]$fitted - boston[(tt+1):TT, 14])^2)), "\n"
}

# Compare with
apply(ff, 2, function(x) { sqrt(mean((x - boston[(tt+1):TT, 14])^2)) })
Description

Combine different forecasts using complete subset regressions. Apart from the simple averaging, weights based on information criteria (AIC, corrected AIC, Hannan Quinn and BIC) or based on the Mallow criterion are also available.

Usage

`Forecast_comb_all(obs, fhat, fhat_new = NULL)`

Arguments

- `obs`: Observed series.
- `fhat`: `fhat` Matrix of available forecasts.
- `fhat_new`: Matrix of available forecasts as a test set. Optional, default to NULL.

Details

OLS forecast combination is based on

\[ obs_t = \text{const} + \sum_{i=1}^{p} w_i \hat{obs}_{it} + e_t, \]

where `obs` is the observed values and \( \hat{obs} \) is the forecast, one out of the \( p \) forecasts available.

The function computes the complete subset regressions. So a matrix of forecasts based on all possible subsets of `fhat` is returned.

Those forecasts can later be cross-sectionally averaged to create a single combined forecast.

Additional weight-vectors which are based on different information criteria are also returned. This is in case the user would like to perform the frequensit version of forecast averaging or based on the Mallows criterion (see references for more details).

Although the function is geared towards forecast averaging, it can be used in any other application as a generic complete subset regression.

Value

`Forecast_comb_all` returns a list that contains the following objects:

- `pred`: Vector of fitted values if `fhat_new` is not NULL or the vector of predictions if `fhat_new` is provided.
- `full_model_crit`: List. The values of information criteria computed based on a full model, the one which includes all available forecasts.
- `aic`: A vector of weights for all possible forecast combinations based on the Akaike’s information criterion.
- `aicc`: A vector of weights for all possible forecast combinations based on the corrected Akaike’s information criterion.
A vector of weights for all possible forecast combinations based on the Bayesian’s information criterion.

A vector of weights for all possible forecast combinations based on the Hannan Quinn’s information criterion.

A vector of weights for all possible forecast combinations based on the Mallow’s information criterion.

Eran Raviv (eeraviv@gmail.com)


library(MASS)
tt <- NROW(Boston)/2
TT <- NROW(Boston)
y <- Boston[1:tt, 14] # dependent variable is columns number 14
# Create two sets of explanatory variables
x1 <- Boston[1:tt, 1:6] # The first 6 explanatory variables
x2 <- Boston[1:tt, 7:13] # The last 6 explanatory variables
# create two forecasts based on the two different x1 and x2
coef1 <- lm(y ~ as.matrix(x1))$coef
coef2 <- lm(y ~ as.matrix(x2))$coef
f1 <- t(coef1 %*% t(cbind(rep(1,tt), Boston[(tt+1):TT, 1:6])))
f2 <- t(coef2 %*% t(cbind(rep(1,tt), Boston[(tt+1):TT, 7:13])))
ff <- cbind(f1, f2)
comb_all <- Forecast_comb_all(obs = Boston[(tt+1):TT, 14], fhat = ff)
# To get the combined forecasts from the all subset regression:
Combined_forecast <- apply(comb_all$pred, 1, mean)
# To get the combined forecasts based on aic criteria for example:
Combined_forecast_aic <- t(comb_all$aic %*% t(comb_all$pred))
Index

Forecast_comb, 1
Forecast_comb_all, 3