

the3PRF.zip

Bryan Kelly and Seth Pruitt

2012-9-13

Copyright: Bryan Kelly and Seth Pruitt

First version date: 2010-12-30

Please cite as:

Kelly, Bryan and Seth Pruitt (2012): “The Three-Pass Regression Filter: A New Approach to Forecasting Using Many Predictors,” *Working Paper, Chicago Booth*.

Contains:

- `estimate3PRF.m`
Estimates the Three-Pass Regression Filter (3PRF): forecasts, forecast errors, rsquare, alpha, and forecast/alpha asymptotic variance estimates; uses regression loops to be robust to NaNs
- `forecast3PRF.m`
Forecasts using the Three-Pass Regression Filter (3PRF): runs regression loops to be robust to NaNs
- `readme.pdf`
This readme file

Requires:

- Matlab Stats toolbox

Usage:

- Place files in path or working directory. Call as functions, according to function documentation.

Function Documentation

• estimate3PRF

```
[forecasts] = estimate3PRF(y,X,Z)
[forecasts] = estimate3PRF(y,X,Z,flag1,value1,...)
[forecasts pointests avarests] = estimate3PRF(y,X,Z,flag1,value1,...)
```

INPUTS

- y : (T x 1) time series of target variable
- X : (T x N) matrix of predictors' time series
- Z : (T x L) matrix of proxies' time series
 - *OR* scalar choice for choice of L automatic proxies

NOTE: The timing between y and X/Z must be adjusted before being passed -- we assume that the t-th row of X is being used to forecast the t-th row of y, and that the t-th row of Z are the proxies used to extract info from the t-th row of X. This means that y (or X/Z) should be properly led (lagged) before being passed.

Flags

- 'type' : 'IS Full' (default), 'OOS Recursive', 'OOS Cross Val' , 'OOS Rolling'
- 'window' : 2-vector such as [<number of obs before forecasted to drop> , <number of obs total to drop>], used by OOS Cross Val (default [0 1])
- 'mintrain' : EITHER (1) scalar choice of smallest number of observations for training sample, used by OOS Recursive (default = round(T/2)); OR (2) 2-vector with smallest number as in (1) and <gap between training estimation window and first forecast constructed> (default = 0). Note: sets number of non-missing observations required for each predictor to be included in the training sample
- 'rollwin' : EITHER 2-vector such as [<number of obs in rolling window> , <number of nonmissing obs required for that predictor to be included in the training sample>] (default [30 20]); OR 3-vector with last element <gap between training estimation window and first forecast constructed> (default = 0).
- 'pls' : 'on' or 'off' (default), runs PLS version which normalizes X and runs 1st and 2nd stages without constants [only if scalar is passed for Z]

OUTPUTS

- forecasts : (T x 1) time series of forecasts
- pointests : structure of point estimates with following fields
 - forecasts : (T x 1) time series of forecasts
 - ferrors : (T x 1) time series of forecast errors
 - rsquare : R2 (can be negative for OOS) when compared to rolling historical mean
 - encnew : (non-nan when OOS Recursive) value of Clark and McCracken's (2001) ENC-NEW statistic
 - rollfore : (T x 1) time series of rolling mean forecasts when OOS; nan otherwise
 - *** THE FOLLOWING ARE COMPUTED FOR 'IS Full' -- otherwise not computed ***
 - alpha : (N x 1) coefficient
- avarests : structure of estimates of asymptotic variances if 'IS Full' with following fields
 - forecasts : (T x 1) estimate of asymp covariance of time t forecast
 - alpha : (N x N) estimate of asymp covariance matrix of alpha

DESCRIPTION: Computes Three-Pass Regression Filter (Bryan Kelly and Seth Pruitt, 'The Three-Pass Regression Filter: A New Approach to Forecasting Using Many Predictors'); see paper for more details.

Please cite as: Kelly, Bryan and Seth Pruitt (2011): ‘‘The Three-Pass Regression Filter: A New Approach to Forecasting Using Many Predictors,’’ Working Paper, Chicago Booth.

Assumes the data are timed as follows:

- $y = (y[h+1], y[h+2], \dots, y[T+h])'$
- $X = (x[1]', x[2]', \dots, x[T]')$

A future realized variable (like the target itself) might be one of the proxies, as well as other variables realized at time $t+h$.

Requires Stats toolbox (nanstd.m, nanmean.m, regress.m)

- forecast3PRF

```
[forecasts] = forecast3PRF(y,X,Z)
[forecasts] = forecast3PRF(y,X,Z,flag1,value1,...)
```

INPUTS

- y : (T x 1) time series of target variable
 - X : (T x N) matrix of predictors' time series
 - Z : (T x L) matrix of proxies' time series
- OR
- scalar choice for choice of L automatic proxies.

NOTE: The timing between y and X/Z must be adjusted before being passed. We assume that the t-th row of X is being used to forecast the t-th row of y, and that the t-th row of Z are the proxies used to extract info from the t-th row of X. This means that y (or X/Z) should be properly led (lagged) before being passed.

Flags

- 'type' : 'IS Full' (default), 'OOS Recursive', 'OOS Cross Val'
- 'window' : 2-vector such as [<number of obs before forecasted to drop> , <number of obs total to drop>], used by OOS Cross Val (default [0 1])
- 'mintrain' : EITHER
 - (1) scalar choice of smallest number of observations for training sample, used by OOS Recursive (default = round(T/2))
 - OR
 - (2) 2-vector with smallest number as in (1) and gap between training estimation window and first forecast constructed (default = 0)
- 'pls' : 'on' or 'off' (default), runs PLS version which runs 1st and 2nd stages without constants [only if scalar is passed for Z]

OUTPUTS

- forecasts : (T x 1) time series of forecasts

DESCRIPTION: Computes Three-Pass Regression Filter (Bryan Kelly and Seth Pruitt, 'The Three-Pass Regression Filter: A New Approach to Forecasting Using Many Predictors'); see paper for more details.

Please cite as: Kelly, Bryan and Seth Pruitt (2011): 'The Three-Pass Regression Filter: A New Approach to Forecasting Using Many Predictors,' Working Paper, Chicago Booth.

Assumes the data are timed as follows:

- y = (y[h+1] ,y[h+2] ,...,y[T+h])'
- X = (x[1]', x[2]', ...,x[T]')'

A future realized variable (like the target itself) might be one of the proxies, as well as other variables realized at time t+h.

Requires Stats toolbox (nanstd.m,nanmean.m,regress.m)