Problem Set 2 ECON8825

Federico Rodari

October 28, 2022

1 Models

AR(p) I decided to adopt an information criterion to chose the optimal lag. The BIC suggests $p = 4$, while the AIC is less conservative by construction.

Factor Model The eigenvalues resulting from the decomposition of $X'X$ decay relatively fast, since the first 20 out of approximately 100 explain almost 100% of the variability in the data. Therefore, I decided to use $r = 20$.

LASSO To choose the optimal λ I performed block cross-validation (`blockcv.m`), selecting $k = 10$ non-random nodes that creates non-overlapping sets with almost the same number of observations.

RIDGE To choose the optimal λ I performed block cross-validation (`blockcv.m`), selecting $k = 10$ non-random nodes that creates non-overlapping sets with almost the same number of observations. I faced some issues when selecting the appropriate grid, as I find out that the optimal λ tends to be (extremely much) higher than the one chosen by the LASSO on average.

Bagged Tree I grow a deep regression tree and use block-boostrap to reduce the variability of the tree estimates. I Initially wanted to implement a proper Random Forest, but I find it to be not very intuitive when relying on Matlab built-in functions (especially when I have to adapt it to a time series structure). Moreover, I constructed a function (`blockbootstrp.m`) that performs block bootstrap, as Matlab cannot handle time series bootstrapping. I "cheated" in the sense that for each horizon $T_0 + t$ I was actually drawing a bootstrap sample from the whole dataset, but the function I created currently works only for a dataset with T even.

Ensemble For this model I am facing a problem. I see the question expects α_j to be time-varying, but I do not get how I can get an estimate for the first periods when I do not have enough degrees of freedom. I cannot augment the dataset with the training data of $UNRATE$ as the columns would be trivially collinear, so I estimated the ensemble assuming the weights were time invariant.

2 Results

Figure 1 shows graphically how the different models perform in predicting the one month ahead unemployment growth rate, while table 1 shows the mean squared error over the test sample. As one could expect, the ensemble method is best in terms of predictive performance as it optimally adjusts the weight based on the relative predictive accuracy of each model. Since the RIDGE is the best of the single models (excluding the ensemble), it received the highest weight when averaging the predictions. I cannot conclude much with respect to the nonlinear ML model as I feel I should work more on implementing the random features selection in the model.

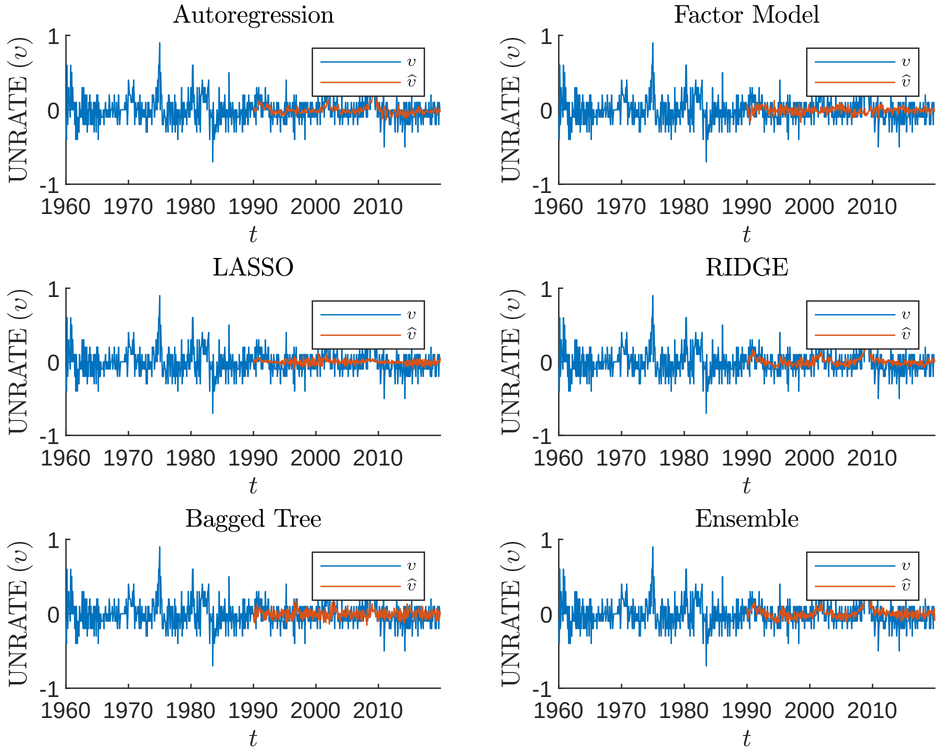

Figure 1: Models performance

Model	MSE
AR(4)	0.0236
Factor Model	0.0270
LASSO	0.0248
RIDGE	0.0202
Bagged Tree	0.0277
Ensemble	0.0181

When comparing predictive performance in expansions and recessions, I partition the time series of prediction errors according to a dummy based on the NBER dating of expansions and recessions (takes value 1 if there is a recession, 0 for expansion) and average the error within each of the two sets. Table 2 shows the results, all models appear to have a better predictive accuracy during recessions, and the models ranking does not change.

Model	Expansion	Recession
AR(4)	0.0244	0.0188
Factor Model	0.0280	0.0205
LASSO	0.0260	0.0172
RIDGE	0.0211	0.0151
Bagged Tree	0.0285	0.0228
Ensemble	0.0185	0.0161

The way I understood the question was to simply evaluate the predictive accuracy we already had from the previous exercises, so I only evaluated the accuracy during the recessions in the test sample (1990-2019).