**Greedy Function Approximation for Macroeconomic Forecasting**

Darren Homrighausen, Daniel J. McDonald, and Cosma Shalizi

Vector autoregression (VAR) is a core tool in macroeconomic forecasting for predicting a multivariate time series $Y_t \in \mathbb{R}^q$ for $t = 1, \ldots, T$. Once the lag parameter $J$ is set, a VAR can be written

$$Y_t = \sum_{j=1}^{J} A_j Y_{t-j} + \epsilon_t,$$

where $A_j \in \mathbb{R}^{q \times q}$ are the parameter matrices and $\mathbb{V}[\epsilon_t] = \Sigma$. VARs are straightforward to fit through the use of least squares. Nonlinear adaptations of VARs have been in use for some time now, such as the single-index econometric model (SIAVAR) [6]

$$Y_t = F(Y_{t-1}, \ldots, Y_{t-J}) + \epsilon_t = \sum_{j=1}^{J} f_j(A_j Y_{t-j}) + \epsilon_t,$$

where $A_j$ are the single-index weights. This model directly generalizes the VAR as it allows for marginally nonlinear functions of $Y_{t-j}$ via $f_j$. The single-index model is a special case of an older, more general method in statistics known as 'projection pursuit' [5]. While the SIAVAR allows for interactions across the time series at a particular time $t$, it does not allow for cross-time interactions. Fully additive models, where $F(Y_{t-1}, \ldots, Y_{t-J}) = \sum_{q=1}^{Q} f_q(Y_{t-1}, \ldots, Y_{t-J})$, generalize projection pursuit and allow for more complicated interactions across both series and time, however, they can be difficult fit numerically due to a nonconvex objective function and identifiability concerns.

One popular method in machine learning for successfully fitting additive models is known as (gradient) boosting [3, 4]. Boosting fits an additive model where $f_q$ is a *base-learner*—commonly a simple method such as marginal regression or heavily-pruned decision trees. The base-learner is refitted on weighted versions of the data such that observations that have large residuals or are misclassified are emphasized in successive fits.

Our research develops tools for fitting additive time series models and examines their utility for forecasting. Our results show that making predictions via boosting with a multivariate tree base-learner improves on VARs in a variety of situations, even when the order of the VAR is chosen in an oracle fashion. Additionally, though there has been some recent work in the area of boosting time series data [1, 2, 7, 8], there are many open avenues of research. First, boosting readily allows for loss functions other than squared error, both for continuous-valued $Y_t$ (such as in robust regression) and for categorical-valued $Y_t$ (such as the multinomial logistic likelihood). Second, the implications of various choices—e.g. the base learner and the number of boosting iterations—are poorly understood in the time series regime, but have important consequences for predictive performance.

Relevance to the 2015 NBER-NSF Time Series Conference

In this paper, we seek to expand the use of boosting in the time-series community by demonstrating its superiority as a forecasting technique. Boosting is a central technique in machine learning applications with independent and identically distributed data but is just beginning to be developed for time-series applications and used by time-series analysts. Therefore, this work is very relevant to the NBER-NSF conference as it is a major venue through which statistics, machine learning, and economics communities interface and exchange ideas. This work is partially funded by a grant from the Institute for New Economic Thinking.

## References

[1] T. Buchen and K. Wohlrabe. Forecasting with many predictors: Is boosting a viable alternative? *Economics Letters*, 113(1):16–18, 2011.

[2] P. Bühlmann and B. Yu. Boosting with the l 2 loss: regression and classification. *Journal of the American Statistical Association*, 98(462):324–339, 2003.

[3] J. Friedman, T. Hastie, R. Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2):337–407, 2000.

[4] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.

[5] J. H. Friedman and W. Stuetzle. Projection pursuit regression. *Journal of the American statistical Association*, 76(376):817–823, 1981.

[6] Y. Li and M. G. Genton. Single-index additive vector autoregressive time series models. *Scandinavian Journal of Statistics*, 36(3):369–388, 2009.

[7] S. Ng. Viewpoint: boosting recessions. *Canadian Journal of Economics/Revue canadienne d'économique*, 47(1):1–34, 2014.

[8] S. Ng and J. Bai. Selecting instrumental variables in a data rich environment. *Journal of Time Series Econometrics*, 1(1), 2009.