University of California San Diego

MATH 181E
Mathematical Statistics - Time Series

A Comparative Analysis Between Direct and Iterated AR Methods For Forecasting Macroeconomic Time Series - Empirical Report by Massimiliano Marcellino, James H. Stock, Mark W. Watson.

March 25, 2023

Author	*Student ID*
Daria Barbour-Brown	A17214152

Contents

Introduction

1 Overview

This report first considers the important distinction between what should be theoretically true of forecasting methods and what may be optimal in practice. The goal of the analysis was to address previous findings from theoretical literature that points to direct forecasting as the more efficient method. The authors' apply iterated forecasts and direct forecasts to a set of 170 monthly U.S macroeconomic time series to decide if the theoretical claims about the two methods hold true in an empirical setting.

1.1 Stationarity

Macroeconomic time series appear to be nonstationary. Stationarity is a crucial concept to time series analysis. We learned that stationarity is the notion of regularity in values over some time shift. A strictly stationary time series will have multivariate distributions in agreement with their time shifted counterparts and having such a series makes it easier to estimate a model by averaging since we attain regularity in the mean and autocorrelation functions. These fundamental understandings allowed me to see why it would be important to approximate stationarity. The theoretical literature that this analysis compares against, studies stationary series so it was important to make this transformation. Additionally, all of the theoretical calculations we performed in class regarding forecasting, were made restricting our series to be stationary models. As the first step in this analysis, the authors' took the first or second difference of each series, X_t to estimate the forecasting model. We know this dth-order difference to be defined by $\nabla^d = (1-B)^d$, where $BX_t = X_{t-1}$. With this estimated forecast, the h-step-ahead forecast of the original series would be calculated based only on values of the series up to the date on which the forecast is made otherwise known as recursive.

1.2 Univariate and Bivariate Series

The paper's aim is to make the comparison between iterated and direct forecasting of both univariate and multivariate series empirically. A univariate time series is made up of single observations recorded over equal time increments whereas a multivariate time series has more than one time series variable where each variable depends on its past values and on other variables. One of the most commonly used methods for dealing with multivariate time series forecasting is known as Vector Autoregression (VAR). This is not a topic we reached in our course content, but through the process of dissecting this report, I found that in Vector Regression each variable written as a linear function of the past values of itself and the past values of all the other variables unlike the AR process (used for univariate time series data), this returns a system of equations or matrix providing the relationships between multiple variable within a series.

Setting Up

2 Notation

For univariate models, X_t denotes the logarithm of the series of interest. This was familiar in our class discussion if linearizing our data by way of logging a series. Forecasts of X_{t+h}, were calculated by information at time t. Then y_t denotes the stationary transformation of the series after taking first or second differences. Specifically, the authors' let X_t be integrated of order d I(d); then, $y_t = \nabla^d X_t$.The order of integration is a new concept for me with regards to time series. With some further research I found that "Order of integration" tells us the minimum number of differences needed to get a stationary series. My understanding of the differencing operator allowed for this to make sense.

2.1 Defining Iterated and Direct Forecasts

Iterated AR Forecast

An iterated h-step-ahead time series forecast is made using a one-step-ahead model that is iterated forward for the desired number of time steps. The iterative method is based on the AR model.

We begin with the One-Step-Ahead AR model for y_t:

$$y_{t+1} = \alpha + \sum_{i=1}^{p} \phi_i y_{t+1-i} + \varepsilon_t.$$

The parameters of the iterated AR forecast are then estimated using Ordinary Least Squares. We know from classical statistics that (OLS) is a type of least squares method for picking unknown parameters in a linear regression. This is done by minimizing the sum of the squares of the differences between the observed and expected values. For the iterated AR model this study constructed a recursive form for the estimated forecast

$$\hat{y}^I_{t+h/t} = \hat{\alpha} + \sum_{i=1}^{p} \hat{\phi}_i \hat{y}^I_{t+h-i/t},$$

In the iterated approach above the forecast is constructed as a function of forecasts at shorter horizons.

Direct Forecasting Regression Model

In the direct approach an h-step-ahead forecast is constructed directly, unreliant on forecasts at shorter horizons. This method was the first introduction to forecasting we were given. The goal is to predict future values and attain a prediction function where the mean square error is minimized. Here we discussed satisfying two prediction equations. This is done by crafting a regression between y_{t+h} and X_t for instance:

$y_{t+h} = a + bX_t + error$, where your forecast for y_{t+h} will be $f(h) = a + bX_t$.

In this method, the estimates for the parameters are minimizing the mean squared error of the h-step-ahead function, again estimated by the OLS regression.

Direct Forecasting Regression Model:

$$y^h_{t+h} = \beta + \sum_{i=1}^{p} \rho_i y_{t+1-i} + \varepsilon_{t+h}.$$

2.2 Lag Order

We were first introduced to the notion of lag at the very beginning on our time-series journey. We know that autocorrelation describes the similarity between a time series and a time-shifted version of the same time series. This shift is known as lag. Calculating autocorrelation is a familiar topic but the detailed process of selecting lag-order in a comprehensive data analysis was new to me. With some further digging I found that choice of lag length is considered to be an empirical issue. With regards to forecasting, a researcher will estimate successive lags but it will come with a cost. As these estimations are made, we lose degrees of freedom in our model. Specifically in economic time series data estimations of lags are correlated to increasing the likelihood of multicollinearity (also a new term, describing a model that has dependency on multiple independent variables). This can drastically complicate your results.

For this study, the data is given monthly and there seems to be a general consensus that for monthly data, 6, 12 or 24 lags can be used so it is interesting that the authors chose 4 as one of their lag-orders to study; likely to round out the study in low-lag conditions. I also found that high lags can give imprecise estimation where low lags can give specification errors. My source suggests that sometimes the best way to choose lag order is via Akaike Criterion or Bayesian Information Criterion which is one of the methods this report choses to implement. “Various simulation studies have tended to verify that BIC does well at getting the correct order in large samples, whereas AIC tends to be superior in smaller samples where the relative number of parameters is large” (Shumway, 50). Based on this notion it makes sense that the analysts wanted to test lags using both types of computation. We learned that in order to measure accuracy of a model we can find SSE for each model. This alone is not enough to measure the goodness of fit, we know that the more parameters we add to a model, the larger the variance, which can cause overfitting. Conversely, minimizing the number of parameters can yield a high bias, leading to an under fitted model. The AIC and BIC computations should allow us to find a balance between the number of parameters and the residuals.

In this paper the authors used two fixed lag-orders, 4 and 12 and two chosen by way of AIC or BIC. For the iterated forecasts, the AIC and BIC were computed using sum of squared residuals (SSR) from the one-step-ahead regression. Where for the direct forecasts, the AIC and BIC were computed using the SSR from the estimated h-step-ahead regression. Again, these 4 choices were based on the theoretical suggestions they intended to empirically test against.

2.3 Mean Square Prediction Error

In this analysis, the recursive forecast error is $e_{t+h} = \hat{X}_{t+h} - X_{t+h}$, and the sample MSFE is

$$\text{MSFE} = \frac{1}{T_2 - T_1 + 1} \sum_{t=T_1}^{T_2} e_{t+h}^2.$$

. Where T_1 denotes the date at which the first residual forecast is made, and T2 denotes the final date. This sample MSFE is computed for each 170 series, for both the iterated method (with 4 lag choices), and the direct method (with 4 lags choices), and for each horizon (3, 6, 12, and 24 months). Finally, for each series and horizon, the empirical efficiency of comparable direct and indirect forecasts was assessed by comparing these respective MSFEs.

Interpretation and Analysis

3 The Data

For this report the analysis looks at 170 major monthly U.S. macroeconomic time.The data was split into five categories of series:

(A) Income, output, sales, and capacity utilization (38 series);

(B) Employment and unemployment (27 series);

(C) Construction, inventories and orders (37 series);

(D) Interest rates and asset prices (33 series);

(E) Nominal prices, wages, and money (35 series).

3.1 Transformations

For a simpler analysis, series that represented quantities, indexes, and price levels were transformed to logarithms and series representing interest rates or unemployment rates did not undergo any transformation. Next the set of series were differenced (integrated of order zero). Separately, “real” quantities and real prices were treated as I(1) and the process was repeated for the series in category (E) I(2), a series describing prices, wages, and money. This process of categorization and manipulation is not something I am super comfortable with yet. I would have liked to see some visualizations of these transformations to better understand why and how order of integration is specifically used to attain consistent analysis because my knowledge of this idea is limited to its ability to estimate stationarity.

3.2 How The Report Compared Results

For univariate autoregressions the analysts took a look at distributions of the ratios between MSFE of direct forecast to the MSFE of the iterated forecast. These comparisons kept lag values consistent and allowed forecasting horizons to vary. Of course we know that the forecast horizon is the length of time into the future for which forecasts are to be made. A mean relative MSFE of 0.99 for example suggested that the direct estimator on average is slightly better than the indirect estimator. The authors also addressed the bootstrap p-values to test the efficiency of the iterated estimator. Bootstrapping is a concept touched on briefly in our text but its overall implementation was not discussed. Bootstrapping mimics the sampling process by assigning measures of accuracy (bias, variance, confidence intervals, prediction error, etc.) to sample estimates. With these values the report finds that a good choice between iterated or direct estimators will depend on the method of lag selection and from the bootstrap p-values, these improvements are typically statistically significant.

Next the authors excluded the series containing price, wage, and money data to find that of the other series, an iterated forecast is preferred to the direct forecast at all horizons. On the other hand, for the price, wage, and money series, the direct estimator proved to be stronger at all horizons. The interesting takeaway here is that it is always a good idea to find outlying sets and then to deal with that data separately so as not to skew overall findings. The full analysis looks at all 4 models; (AR(4), AR(12), BIC and AIC) for each of the 5 data categories, comparing both MSFE ratios and forecasting horizons.

A similar analysis was done for the bivariate models through the use of stratified random subsample of the models. 2000 samples were randomly drawn from these subsamples to keep the data set at a more reasonable size. This area of the report was harder for me to follow but the strategies used for bivariate time series analysis are really interesting! Essentially, the authors needed to compare a total of 170 series all against each other, giving a grand total of 28,730 possible pairs of bivariate data (VARs). Since the series are organized by the above five categories, the analysis could make 10 pairs of non repeated categories. Then sampling 200 pairs from each of the 10 paired categories gave a new total of 2000 series, a much more controllable set. Similar comparisons were made as with the univariate series with similar conclusions.

4 Concluding Results of The Study

From a large data set of monthly U.S. macroeconomic time series this report found that iterated forecasts tend to have smaller MSFEs than direct forecasts while lag-length choice places a big role in this relative improvement over a direct forecast. On the other hand, direct forecasts seem to get worse with larger forecasting horizons suggesting that iterated forecasts can give us pretty good approximations to the best linear predictor. Overall the iterated forecasts outperform the direct forecasts, a peculiar conclusion that contradicts much of the reports supported theoretical literature.

Conclusion

5 Final Remarks

Perhaps one of the biggest takeaways I got from the report was the importance of bias/variance tradeoff. Finally I feel as though I understand how minimizing these errors help statisticians to avoid over and under-fitting in their predictive models. In terms of forecasting bias is defined to be the difference between the true value we mean to predict, and the average value our model provides, which variance gives us the spread in our data. Our goal is to create models that balance these errors.

Overall I was surprised to find that dissecting this report did not feel too uncomfortable. Much of the concepts were familiar, and it was pleasant to revisit ideas about stationarity, Auto Regressive Models, iterative forecasting, bias and variance, as well as direct prediction. Upon first glance, I had trouble relating the two methods of forecasting to the specific theoretical computations we did in class but in fact reading through the report provided me with the bigger picture I needed to understand their differences in application.

References

ADELEYE, Bosede Ngozi. "Time Series Analysis (Lecture 2): Choosing Optimal Lags in Eviews." CrunchEconometrix, 1970, http://cruncheconometrix.blogspot.com/2018/02/time-series-analysis-lecture-2-choosing.html.

Shumway, Robert H., and David S. Stoffer. Time Series Analysis and Its Applications: With R Examples. Springer, 2017.

Singh, Seema. "Understanding the Bias-Variance Tradeoff." Medium, Towards Data Science, 9 Oct. 2018, https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229.

Stephanie. "Order of Integration: Time Series and Integration." Statistics How To, 4 Mar. 2023, https://www.statisticshowto.com/order-of-integration/.