Forecasting

Bernt Arne Ødegaard

24 November 2021

Contents

1 Forecasting

The classical way of thinking about forecasting is that it is the estabishment of a *probability* statement about the future value of an economic variable.[1]

If we let x_t be the variable of interest, we want to evaluate the conditional distribution of x_{t+1} based on a set of data observed *prior* to date $t+1$. The choice of what data to use to estimate the conditional distribution is up to the econometrician. It can be past values of the same variable $(x_t, x_{t-1}, \ldots)$, or past values of other variables $(\mathbf{y}_t, \mathbf{y}_{t=1}, \ldots)$.

To do inference it is necessary that we are asking a meaningful question, that the relevant distributions exist.

Of particular interest is the *point estimate* of the variable x_{t+1}.

The best estimate of this is the conditional expectation, conditioned on all available information at time t.

In this course we are not going to go through all relevant issues in forecasting, we will limit the discussion to some common issues that you are likely to run into in finance.

In this introduction we will comment on some other issues, though.

1.1 Choice of forecasting model - theory

If one want to have a theoretical basis for a choice of a forecasting model, it can not be divorced from the question of *what is the forecast used for*?

The relevant theoretical framework for choosing a forecasting model is *decision theory*: The forecast is used to change the economic decisions of some decision maker. One want to choose

[1] Haavelmo (1944).

the forecasting model to maximize the utility of this decision maker. See Granger and Machina (2006). See also Geweke and Whiteman (2006) for the Bayesian perspective.
So keep in mind that this kind of thinking is ideal.

1.2 Choice of forecasting model - common practice

In practice the modelling choices tend to be more *ad hoc*.
One specifies a parsimonious (which typically will be linear) specification in the variables of interest, and then uses a number of econometric techniques to evaluate them.
This is the approach we take here: Illustrate the econometric techniques used to evaluate forecasting models.

- In sample testing of model fit.
  - Nested models
  - Non-nested models
- In sample forecast evaluation.
- Out of sample forecast evaluation.

Use a simple forecasting exercise (from Næs, Skjeltorp, and Ødegaard (2011)) to illustrate the issues.
Let us first go through this abstractly before getting to specifics of the model

1.3 In sample testing of model fit.

Let x_t be the variable we want to forecast, and y_t some other variable.
Write down the linear ADL(1,1) model

$$x_{t+1} = a + b_x x_t + b_y y_t + \varepsilon_{t+1}$$

A typical question to ask in this context: Does y_t have information about the future x_{t+1} (or – is it useful for forecasting).
Answering this: a standard test of whether b_y is significant.
So we can use standard methods here.
What about if we have two competing explanatory variables: y_{t-1} and z_{t-1}, and we want to compare the two specifications:

$$x_{t+1} = a + b_x x_t + b_y y_t + \varepsilon_{t+1}$$

$$x_{t+1} = a + b_x x_t + b_z z_t + \varepsilon_{t+1}$$

Is this a non-nested test?
Not really, we can write a model that nests both specifications

$$x_{t+1} = a + b_x x_t + b_y y_t + b_z z_t + \varepsilon_{t+1}$$

and test restrictions on the coefficients b_y and b_z.
A better example of non-nested specifications is AR specifications vs MA specifications.

1.4 In sample forecast evaluation

A forecast evaluation is based on comparison of a forecast with the realized value.
Presumably, the better the forecast, the closer the forecast is to the realization.
Compare two forecasting models, one with, and one without y as explanatory variable:

$$\hat{x}^a_{t+1} = a + b_x x_t + \varepsilon_{t+1}$$

$$\hat{x}^b_{t+1} = a + b_x x_t + b_y y_t + \varepsilon_{t+1}$$

Forecast evaluation is based on the difference between the forecasts and realizations

$$e^a = x_{t+1} - \hat{x}^a_{t+1}$$

$$e^b = x_{t+1} - \hat{x}^b_{t+1}$$

Need a metric to aggregate this over many forecasts. Common example: Mean Squared Error (But there are many alternatives, see later)

$$MSE = \sum_t (x_{t+1} - \hat{x}_{t+1})^2$$

Note that in sample the MSE is the same as the squared sum of residuals (in sample).
An equivalent way of comparing two specifications like the above: Difference in R^2 (often presented as the increase in R^2 when adding one explanatory variable.

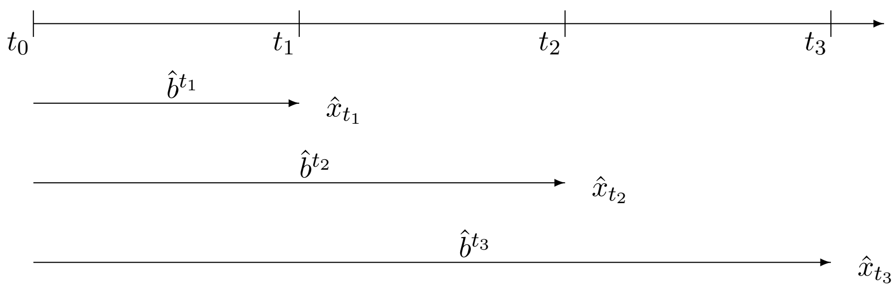
1.5 Out of sample forecast evaluation

In the above, the whole data series is used to construct parameter estimates.
You use *future* information to estimate parameters.
Is this a fair method to evaluate forecasts?
The *out of sample* method is an attempt to get around such critiques. Here, one is careful to only use *past information* to construct any forecast.
So, in the comparison of two linear specifications

$$\hat{x}^a_{t+1} = a + b_x x_t + \varepsilon_{t+1}$$

$$\hat{x}^b_{t+1} = a + b_x x_t + b_y y_t + \varepsilon_{t+1}$$

One adopts a rolling forward procedure



Each forecast is a function of a different parameter $\hat{b}^{t_i}$, $i = 1, 2, \ldots$, a parameter that is re-estimated based on information up to time t_i.
One again constructs the difference between the forecast and realization, and aggregate.
One can also compare mean squared differences, but here it is more common to use a *out-of-sample forecast evaluation criterion.*
We will mention some common choices in the more recent literature.

1. The Diebold and Mariano (1995) is calculated as follows: Suppose we have a candidate predictor i and a competing predictor k. We want to test the null hypothesis of equal predictive accuracy that $E[\bar{d}] = 0 \,\forall\, t$, where $\bar{d} = P^{-1} \cdot \sum_t (\varepsilon^2_{k,t+1} - \varepsilon^2_{i,t+1})$, P is the number of rolling out-of-sample forecasts, and $\varepsilon^2_{k,t+1}$ and $\varepsilon^2_{i,t+1}$ are the squared forecast errors from the two models. The DM statistic is calculated as

$$DM = \frac{\bar{d}}{(\sigma^2_{\bar{d}}/P)^{1/2}}, \tag{1}$$

2. The encompassing test (ENC-NEW) is proposed by Clark and McCracken (2001) The ENC-NEW test asks whether the restricted model (the model that does not include y) encompasses the unrestricted model that includes y. If the restricted model *does not* encompass the unrestricted model, that would mean that the additional predictor (y) in the larger, unrestricted model improves forecast accuracy relative to the baseline. Clark and McCracken show that the ENC-NEW test has greater power than tests for equality of MSE. The ENC-NEW test statistic is given as

$$\text{ENC-NEW} = (P - h + 1) \cdot \frac{P^{-1} \sum_t \left[\varepsilon^2_{r,t+1} - \varepsilon_{r,t+1} \cdot \varepsilon_{u,t+1}\right]}{MSE_u}, \tag{2}$$

where P is the number of out-of-sample forecasts, $\varepsilon_{r,t+1}$ denotes the rolling out-of-sample errors from the restricted (baseline) model that excludes y, $\varepsilon_{u,t+1}$ is the rolling out-of-sample forecast errors from the unrestricted model that includes y, and MSE_u denotes the mean squared error of the unrestricted model that includes y.

3. An F-type test for equal MSE between two nested models proposed by McCracken (2007), termed MSE-F. This test is given by

$$\text{MSE-F} = (P - h + 1) \cdot \frac{MSE_r - MSE_u}{MSE_u}, \tag{3}$$

where MSE_r is the mean squared forecast error from the restricted model that excludes y, and MSE_u is the mean squared forecast error of the unrestricted model that includes y. Both the ENC-NEW and MSE-F statistics are nonstandard and one need to use the bootstrapped critical values provided by Clark and McCracken (2001).

These are illustrative that out-of-sample evaluation tend to get more involved than in-sample evaluation, but are clearly closer to the actual *use* of forecasts.

2 Using R for forecasting

We will be using R to illustrate usage.
Let us first comment on some issues with using R for forecasting.

2.1 Data Handling

The data tend to be time series, for which one need to use a library for time series data: Possibilities

- `ts`
- `zoo` / `zooreg`
- `xts`

2.2 Libraries

A number of different librarires are relevant

2.2.1 Time series methods

- Specifying dynamic linear models: `dyn` / `dynlm`
- Vector Auto Regressions: `vars`

2.2.2 Specialized forecasting library

R has the `forecast` library, which has most of what you need for univariate forecasting, but there are also tests for multivariate cases, such as the diebold mariano test.
Will look at examples of usage.

3 Metrics for comparing forecasts

As mentioned, there are numerous possible metrics for comparison of forecast.
Need a metric for asking "how close" the forecasts are to the realizations.
Common examples:
Mean Squared Error

$$MSE = \frac{1}{T-(T_1-1)}\sum_{t=1}^{T}(y_{t+s}-f_{t,s})^2$$

Mean Absolute Error

$$MAE = \frac{1}{T-(T_1-1)}\sum_{t=1}^{T}|y_{t+s}-f_{t,s}|$$

Mean Absolute Percentage Error

$$MAPE = \frac{1}{T-(T_1-1)}\sum_{t=1}^{T}\left|\frac{y_{t+s}-f_{t,s}}{y_{t+s}}\right|$$

Adjusted AMAPE

$$MAPE = \frac{1}{T-(T_1-1)}\sum_{t=1}^{T}\left|\frac{y_{t+s}-f_{t,s}}{y_{t+s}+f_{t,s}}\right|$$

Theils U-statistic

$$U = \sqrt{\sum_{t=T_1}^{T}\frac{\left(\frac{y_{t+s}-f_{t,s}}{x_{t+s}}\right)^2}{\left(\frac{y_{t+s}-fb_{t,s}}{x_{t+s}}\right)^2}}$$

where fb is a *benchmark* forecast.
Alternative, closer to economic penalty function:
Count number of successful predictions of right sign.
Test for whether you can do better than pure chance.

Literature (on forecast evaluation) The state of the art in terms of forecast evaluation is summarized in Clark and McCracken (2011), which again builds on West (2006)

4 Readings

Elliot and Timmermann (2008)

References

Todd E Clark and Michael W McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105:85–110, 2001.

Todd E Clark and Michael W McCracken. Advances in forecast evaluation. WP 11/20, Federal Reserve Bank of Cleveland, September 2011.

Francis X Diebold and Roberto S Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, 1995.

Graham Elliot and Allan Timmermann. Economic forecasting. *Journal of Economic Literature*, 46(1):3–56, 2008.

John Geweke and Charles Whiteman. Bayesian forecasting. In Graham Elliott, Clive W J Granger, and Alan Timmermann, editors, *Handbook of Economic Forecasting*, volume 1, chapter 1, pages 1–80. Elsevier, 2006.

Clive W J Granger and Mark J Machina. Forecasting and decision theory. In Graham Elliott, Clive W J Granger, and Alan Timmermann, editors, *Handbook of Economic Forecasting*, volume 1, chapter 2, pages 81–98. Elsevier, 2006.

Trygve Haavelmo. The probability approach to econometrics. *Econometrica*, 12:1–118, supplement 1944.

Michael W McCracken. Asymptotics for out-of-sample tests for Granger causality. *Journal of Econometrics*, 140: 719–752, 2007.

Randi Næs, Johannes A Skjeltorp, and Bernt Arne Ødegaard. Stock market liquidity and the Business Cycle. *Journal of Finance*, LXVI:139–176, February 2011.

Kenneth D West. Forecast evaluation. In G Elliot, C W J Granger, and A Timerman, editors, *Handbook of Economic Forecasting*. North-Holland, 2006.