Forecasting and measuring forecast accuracy

Susan Thomas

October 6, 2009

Goals

- Recap: Forecasting from an AR/MA/ARMA process
- Measuring forecast accuracy.

Recap: Principles of time-series forecasting

- x_t is a stochastic process with date from $t = 1, \ldots, T$.
- We want to forecast x_{t+s} for $s = 1, \ldots, S$.
  $s = 1$ means one-step ahead forecasting.
  $s > 1$ means multi-step ahead forecasting.
- Assume that x_t is a stationary process, forecasting uses
  - values of the variable till T.
  - Assumptions about the properties of the process.
- Advantage of time series models (over causal models): easy to compute, becomes a benchmark/preliminary step for further modelling.
- 1970's: controversies about forecasting performance of ARIMA vs. causal models – ARIMA models had better forecast performance.

Principles of forecasting

- All AR/MA/ARMA model can be expressed in three different ways:
  1. The difference equation: $(1 - \phi L)x_t = (1 - \theta L)\epsilon_t$
  2. In past innovations: $x_t = (1 - \hat{\theta}(\infty)L)\epsilon_t$
  3. In past values: $x_t = \hat{\phi}(\infty)Lx_t + \epsilon_t$
- Each ought to give the same answer but makes different demands for information for forecasting.
- Easiest to compute: **the difference equation**

Some observations about the forecast errors

- Notation: $\hat{x}_t(l)$ is the forecast at time t for lead time l, or l periods ahead. This is the minimum square error (mse) forecast for x_{t+l}.
- $\hat{x}_t(l) = E(x_{t+l}|I_t)$, the conditional expectation of x_{t+l}.
- One-step ahead forecast errors are uncorrelated.

$$\hat{\epsilon}_t(1) \text{ vs. } \hat{\epsilon}_{t+1}(1)$$

- Multi-step ahead forecast errors are correlated.

$$\hat{\epsilon}_t(1) \text{ vs. } \hat{\epsilon}_t(2)$$

Inputs to calculate a multi-step ahead forecast for an ARMA DGP

- If DGP is
$$x_t = \phi(q)Lx_t + \epsilon_t + \theta(q)L\epsilon_t$$
,
the forecast $\hat{x}_t(l)$ requires:
  - The x_{t-j} terms that have already happened at the origin (t) are left unchanged.
  - The x_{t+j} terms that have not happened at t are replaced by their forecasts at t, $\hat{x}_{t+j}$.
  - The ϵ_{t-j} which have happened at t are replaced by their model forecast errors $x_{t-j} - \hat{x}_{t-j}$.
  - The ϵ_{t+j} which have not happened at t are replaced by 0.
- Once this is available, forecasting by the difference equation can be done.

Recap: Forecasting using an AR model

One–step ahead forecast for y_t given $I_t = (y_{t-1}, y_{t-2}, \ldots)$

- If the model is AR(1), one-step ahead forecast, $x_t(1)$:

$$E(y_{t+1}|I_t) = \hat{y}_{t+1} = a + \phi_1 y_{t-1}$$

- The one-step ahead forecast MSE of an AR(1) model

$$\begin{aligned} E(y_t - \hat{y}_t)^2 &= (a + \phi y_{t-1} + \epsilon_t - a - \phi y_{t-1})^2 \\ &= \sigma^2 \end{aligned}$$

Recap: Multi-step ahead forecasts for an AR(1) model

- MSE of a two-step ahead forecast for AR(1)

$$
\begin{aligned}
E(y_{t+1} - \hat{y}_{t+1})^2 &= (a + \phi y_t + \epsilon_{t+1} - a - \phi \hat{y}_t)^2 \\
&= (\phi(a + \phi y_{t-1} + \epsilon_t) + \epsilon_{t+1} - \phi(a - \phi y_{t-1}))^2 \\
&= (\phi \epsilon_t + \epsilon_{t+1})^2 \\
&= (1 + \phi^2)\sigma^2
\end{aligned}
$$

- The MSE for an *s*-step ahead forecast for AR(1) is:

$$E(y_{t+s} - \hat{y}_{t+s})^2 = (1 + \phi^2 + \phi^4 + \phi^6 + \ldots + \phi^{2(s-1)})\sigma^2$$

- The MSE becomes larger for longer forecasting horizons, and is $\sigma^2/(1 - \phi^2)$ asymptotically.

Recap: Forecasting using an MA model

- **One–step ahead forecasts**: We have $I_t = (y_{t-1}, y_{t-2}, \ldots, \epsilon_{t-1}, \epsilon_{t-2}, \ldots)$
  We need to forecast y_t.
- If the model is MA(1),

$$\begin{aligned}\hat{y}_t &= \alpha + E(\epsilon_t) + \theta_1 \epsilon_{t-1} \\ &= \alpha + \theta_1 \epsilon_{t-1}\end{aligned}$$

- Forecast error: Mean Squared Error (MSE).

$$\begin{aligned}\text{MSE} = (y_t - E(y_t | I_t))^2 &= (\epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \ldots) \\ &\quad -(\theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \theta_3 \epsilon_{t-3} + \ldots) \\ &= \sigma^2\end{aligned}$$

Measuring forecast accuracy

Features of the forecast errors

- There are several methods to measure the accuracy of the forecasting model – all of them are based on the behaviour of the *forecast errors*.
- Here, there are two kinds of errors:
  1. In-sample errors.
  2. Out-of-sample errors.
- If the model parameters is such that it minimises the forecast errors, the "'one-step-ahead" forecast errors, $x_t(1) - \hat{x}_t(1)$ are:
  - On average, these should be zero.
  - On average, these should be uncorrelated.
- These features must be true for both the "in-sample" and "out-of-sample" forecast errors.
  This becomes a method to select between alternative models for the DGP.

Forecasting accuracy

- There are diffferent ways to summarise the forecasting accuracy.
  Typically the measure used is the RMSE. However, this
- A crucial ingredient to choosing a evaluation criteria is the purpose for which the forecast is being made.
  There is no one best criteria.
- These are typically to the "out-of-sample" forecast errors for the evaluation of alternative forecasting models.

Forecasting accuracy criteria

Mean absolute deviation/error (MAD/MAE) average of the absolute values of the forecast errors.

Mean absolute percentage error (MAPE)

Correlation of forecasts with actual values regression R^2 of changes in actual values of the variable with the changes in forecasted values.

Percentage of "turning points" forecast typically calculated for a binary variable. (Typical measure used – Kuiper's score – percentage of ones correctly forecasted less zeros incorrectly forecasted.)

Conditional efficiency of two different forecasts $\hat{x}_a, \hat{x}_b$. In a regression of $\hat{x}_a, \hat{x}_b$ on x_t, check if the coefficient on $\hat{x}_b$ is zero.

HW – due on Friday, 9^{th} October

1. For the data series in `samples.rda`, evaluate the out-of-sample “forecast accuracy” of the two “best” forecasting models, using
    1. MAD
    2. RMSE
    3. Correlation of forecasts with actual values.
       (Refer to Granger-Newbold(1973) *Some comments on the evaluation of economic forecasts* in **Applied Economics**, 5, 35–47.)

Some references

1. **Armstrong, J. S. et al** (1978), *Symposium on Forecasting with Econometric Methods*, Journal of Business 51, 547–600.
2. **Armstrong, J. S.** (2001), *Principles of forecasting; A handbook for researchers and practitioners*, Norwell, MA: Kluwer.
3. **Armstrong, J. S.** (2006), *Findings from evidence-based forecasting: methods for reducing forecast error*, International Journal of Forecasting 22, 583–598.
4. **Diebold, F. X. and R. S. Mariano** (1995), *Comparing Predictive Accuracy*, Journal of Business and Economic Statistics 13, 253–263.

Some references

1. **Granger, C. W. J and P. Newbold** (1986), *Forecasting Economic Time Series*, 2nd ed., pg: 276-287, London: Academic Press.
2. **Granger, C. W. J and H. Pesaran** (2000), *Economic and Statistical measures of forecast accuracy*, Journal of Forecasting, 19, 537–560.
3. **Mahmoud, E** (1984), *Accuracy in Forecasting: A survey*, Journal of Forecasting, 3, 139–159.