Introduction to Time Series

Daniel Cullen

December 1, 2017

1 Forecasting

Forecasting is different from estimating causal effects, models that are useful for forecasting do not need a causal interpretation. Measures of fit, such as adjusted R^2 are not very useful when estimating causal effects but are informative about the quality of a forecasting model.

A common approach for modeling univariate time series is the autoregressive (AR(p)) model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + u_t.$$

The value of p is called the order of the AR model. The simplest autoregressive model is the AR(1) model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + u_t.$$

The forecast in the next period based on the AR(1) is given by:

$$\widehat{Y}_{T+1|T} = \widehat{\beta}_0 + \widehat{\beta}_1 Y_T.$$

The forecast error is the difference between the forecasted value and the actual value:

$$\text{Forecast error } = Y_{T+1} - \widehat{Y}_{T+1|T}$$

1.1 n-step-ahead forecasting

Once the parameters of the AR model are estimated, we can forecast an arbitrary number of periods into the future. A 1-step-ahead forecast uses the data we observe to predict one period into the future. So to predict Y_{t+1} the first value we do not observe, we calculate

$$\widehat{Y}_{t+1} = \widehat{\beta}_0 + \widehat{\beta}_1 Y_t.$$

To predict further into the future we must use forecasted values of Y in the forecast. To forecast a 2-step-ahead forecast we use the value we estimated from the 1-step-ahead forecast. So to predict Y_{t+2}, we calculate

$$\widehat{Y}_{t+2} = \widehat{\beta}_0 + \widehat{\beta}_1 \widehat{Y}_{t+1}.$$

This method adds uncertainty to the model that causal inference does not have. Now we have uncertainty about the accuracy of the forecasted values that are used as lagged values in the right hand side, uncertainty about the values of the autoregressive coefficients, and uncertainty about value of the error term in the period being predicted. These uncertainties can be quantified and combined to give a confidence interval for the n-step-ahead predictions. This confidence interval becomes wides as n increases because the number of estimated values used is increasing.

1.2 Confidence Intervals for Forecasting

For a one-step-ahead forecast there are only two sources of error:

- The error because future values of u_t are unknown
- The error in estimating the coefficients.

One-step-ahead forecasts do not have error from forecasted values being included in the right hand side because only observed values are used. To find the confidence interval for a forecast we need to calculate the variance of the forecast error.

$$
\begin{aligned}
var(Y_{t+1} - \widehat{Y}_{t+1}) &= var(Y_{t+1}) + var(\widehat{Y}_{t+1}) \\
&= \sigma_Y^2 + var(\widehat{\beta}_0) + Y_t^2 var(\widehat{\beta}_1) + 2Y_t cov(\widehat{\beta}_0, \widehat{\beta}_1)
\end{aligned}
$$

So, the confidence interval for the forecast is

$$
\widehat{Y}_{t+1} \pm c.v. \sqrt{var(Y_{t+1} - \widehat{Y}_{t+1})}
$$