HA 8.8

Jean Jimenez

HA 8.8

Consider `austa`, the total international visitors to Australia (in millions) for the period 1980-2015.

A

Use `auto.arima()` to find an appropriate ARIMA model. What model was selected. Check that the residuals look like white noise. Plot forecasts for the next 10 periods.

Work

The `auto.arima()` function returns the best ARIMA model for a univariate time series. Here, I pass the `austa` time series object through the function.

Series: austa ARIMA(0,1,1) with drift

Coefficients: ma1 drift 0.3006 0.1735 s.e. 0.1647 0.0390

sigma^2 = 0.03376: log likelihood = 10.62 AIC=-15.24 AICc=-14.46 BIC=-10.57

The function suggests an ARIMA(0,1,1) model with drift. It has the following statistics:

- **Coefficients:**
    - MA1=0.3006 = 0.3006
    - Drift=0.1735 = 0.1735
- **Standard Errors:**
    - MA1=0.1647 = 0.1647
    - Drift=0.0390 = 0.0390

MA1 is the moving average of the model.

Now, I need to check if the residuals look like white noise. I do this by passing our model through the `checkresiduals()` function. This function performs a Ljung-Box test, as well as plots for the residuals.

Residuals from ARIMA(0,1,1) with drift
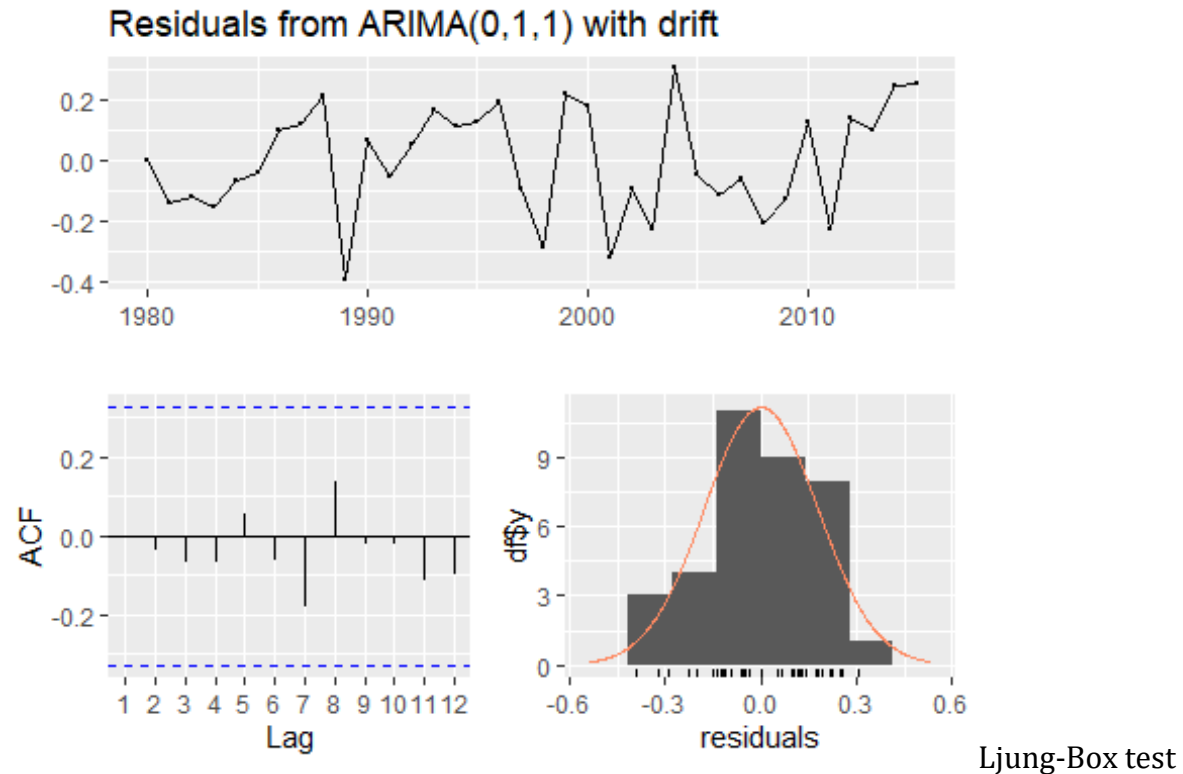
data: Residuals from ARIMA(0,1,1) with drift Q* = 2.297, df = 6, p-value = 0.8905

Model df: 1. Total lags used: 7

The Ljung-box test returned a p-value of 0.8905. This means that our model has no autocorrelation. This test follows the following assumptions, and we failed to reject the null:

- **Null Hypothesis (H0):** The residuals are independently distributed (i.e., there is no autocorrelation).

- **Alternative Hypothesis (H1):** The residuals are not independently distributed (i.e., there is autocorrelation).

The residuals appear to be normally distributed. The ACF plot shows no apparent pattern, and is not significant. This means that it is consistent with white noise and there is no autocorrelation. The time series of the residuals fluctuate around zero with no patterns.

Now, I will plot the forecast for the next 10 periods. For that, I use the `forecast()` function to create the forecast and `autoplot()` to plot:

Forecasts from auto.arima() model
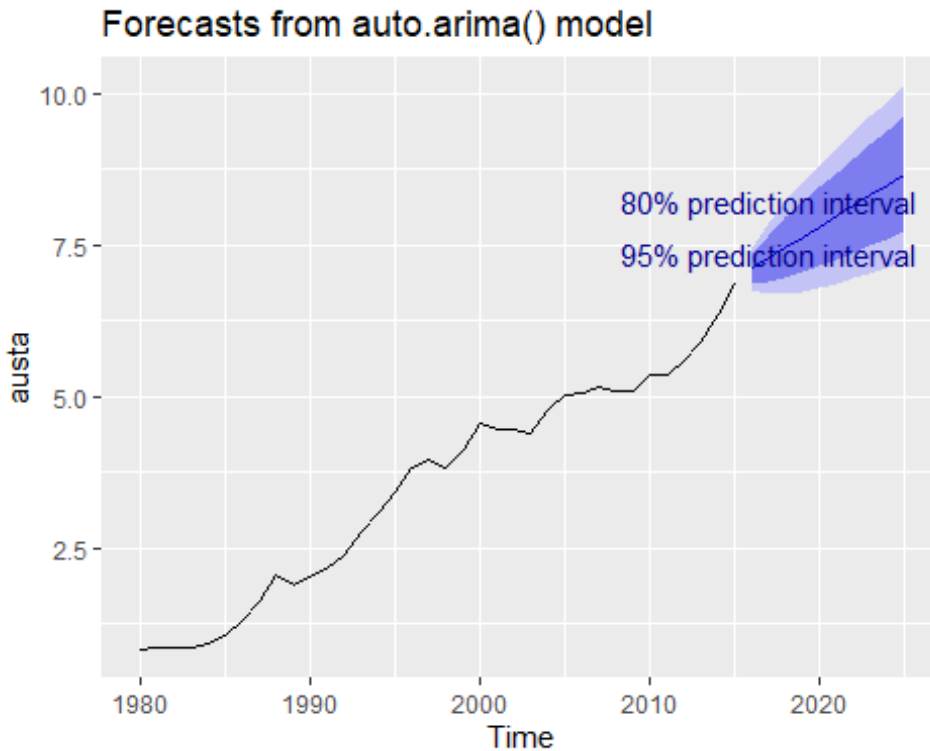
The plot shows the forcasted values for austa, with the 95% prediction interval in dark blue and the 80% prediction interval in light blue.

B

Plot forecasts from an ARIMA(0,1,1) model with no drift and compare these to part a. Remove the MA term and plot again.

Work

Now I have to forecast using ARIMA(0,1,1) without any drift and compare it to the previous forecast. To remove the drift, I set the argument `include.drift=FALSE`.

In the `Arima()` function, `order` follows the following structure: First term is the number of autoregressive terms, second term being the number of differencing terms, and the third being the moving average term.

Forecasts from ARIMA(0,1,1) model with no drift
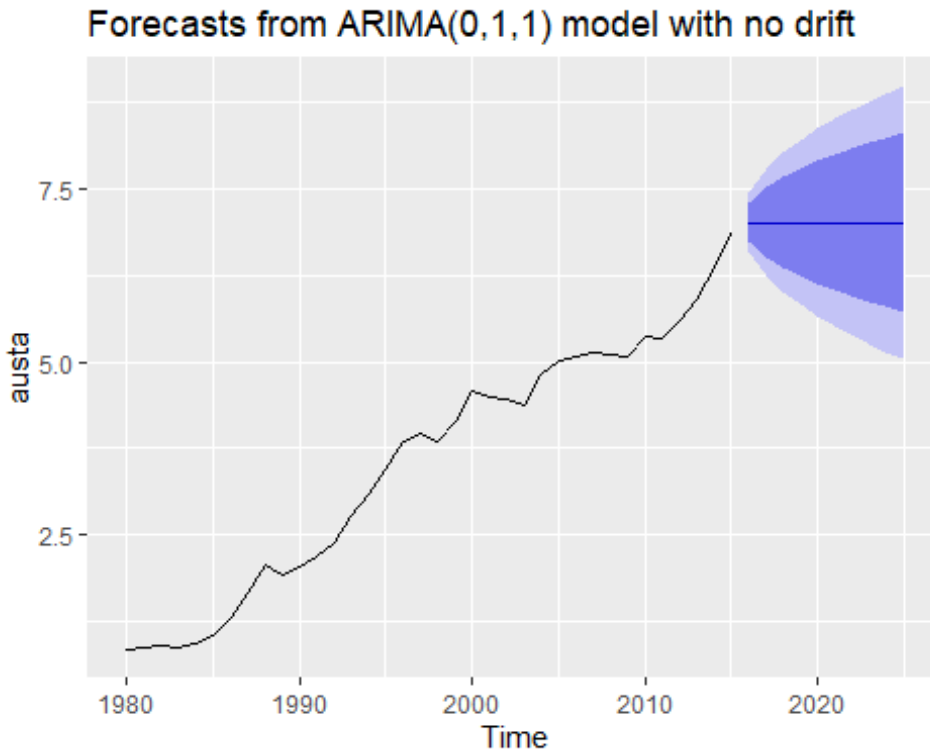
ARIMA(0,1,1) with no drift flat-lines (unlike with drift). This makes sense because without drift removes the difference of the previous data, which will remove the increasing trend.

C

Plot forecasts from an ARIMA(2,1,3) model with drift. Remove the constant and see what happens.

Work

The following is the plot of forecasted ARIMA (2,1,3) with drift.

Forecasts from ARIMA(2,1,3) model with drift
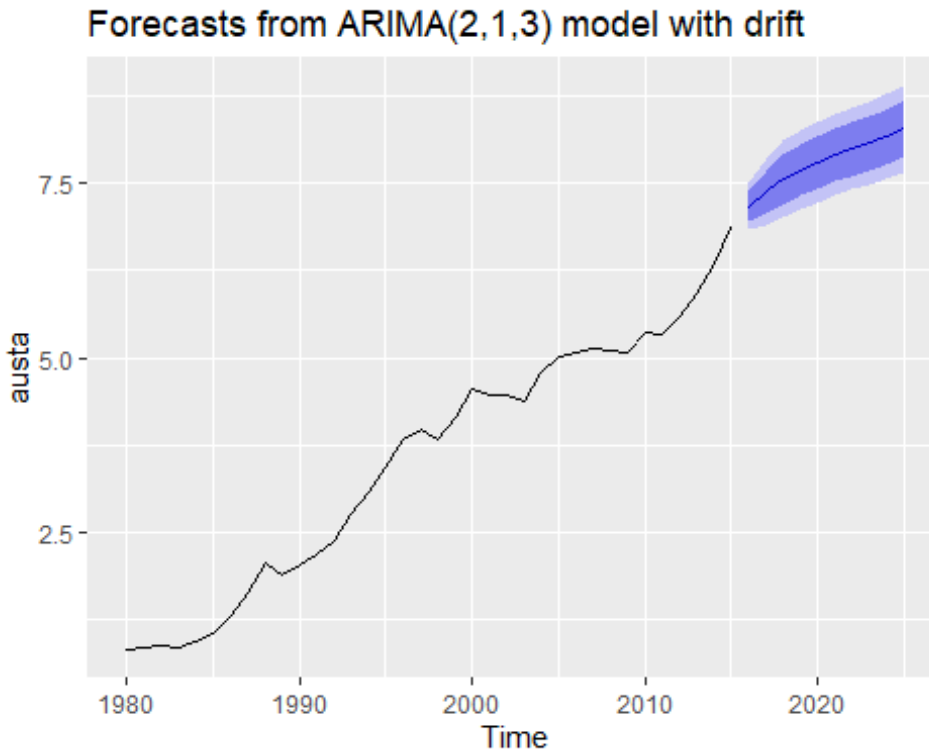
This `ARIMA(2,1,3)` model gives us a forecast of upward trend. It is more reasonable and conservative than the plot in A, but has an increasing trend unlike the plot in B. This is probably because it uses more autoregressive and moving average terms compared to the other two.

Now I will remove the constants from this ARIMA model:

```
tryCatch({
  c_mod2 = Arima(austa, order=c(2,1,3), include.constant=FALSE)
  c_forecast2 = forecast(c_mod2, h=10)
  autoplot(c_forecast2) + ggtitle("Forecasts from ARIMA(2,1,3) model without
constant")
}, error = function(e) {
  print("ARIMA(2,1,3) without constant failed, trying a simpler model.")
  c_mod3 = Arima(austa, order=c(1,1,2), include.constant=FALSE)
  c_forecast3 = forecast(c_mod3, h=10)
  autoplot(c_forecast3) + ggtitle("Forecasts from ARIMA(1,1,2) model without
constant")
})
```

Whenever I try running the ARIMA(2,1,3), the model gives an error and fails. After further research, I have learned that this ARIMA model is not appropriate for this data. This is because there are too many differencing terms. We need to try with a simpler model

D

Plot forecasts from an ARIMA(0,0,1) model with a constant. Remove the MA term and plot again.

The following is the ARIMA(0,0,1) with the constant:

Forecasts from ARIMA(0,0,1) model with no constant
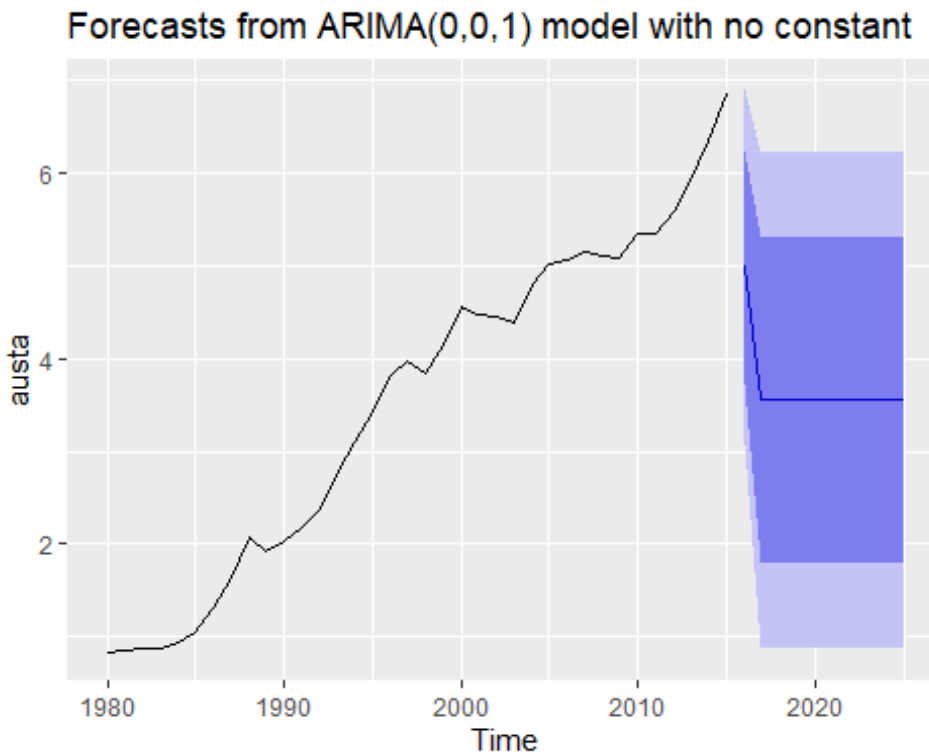
This pattern makes sense. Without a constant or drift, the model does not account for any trends or levels in the data. There is also now higher variability.

E

Plot forecasts from an ARIMA(0,2,1) model with no constant.

Now, I will plot the ARIMA(0,2,1) with no constant. Here we increase the middle term the differencing term by two:

Forecasts from ARIMA(0,2,1) model with no constant
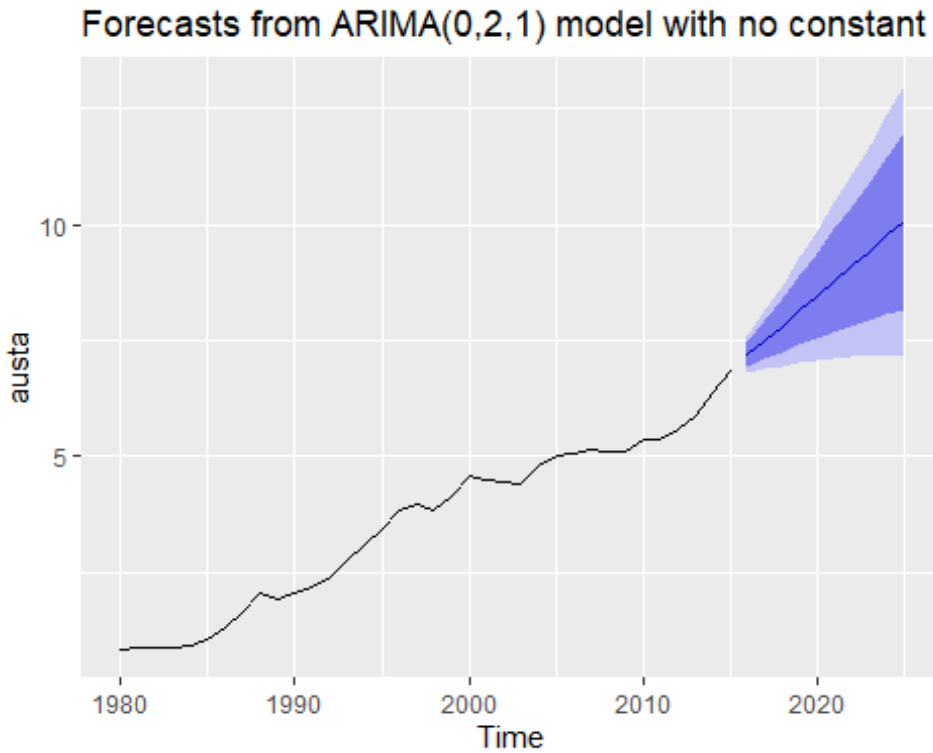
This ARIMA model with two differencing terms accounts for and extrapolates the trend, resulting in an upward-sloping forecast with narrower prediction intervals. By differencing twice, the model effectively removes long-term trends and stabilizes the mean of the series. There is an upward trend in this model.