Hyndman and Athanasopoulos – chapter 9 answers

1. Analysis of *advert* dataset.
   1. Because we can visualize the concomitant changes in advertising spending and car sales.
   2. Done in R. The coefficient of 13.04 suggests that for each extra dollar spent on advertising we expect auto sales to increase by 0.53.
   3. Residuals are autocorrelated. The ACF shows that there are significant autocorrelations up to lag 2; the Breusch-Godfrey test rejects the H0 that there is no autocorrelation (p-value = 0.02)
   4. The standard errors are different, but the estimates of the regression parameters are equal. With an ARIMA (0,0,0) model we are basically assuming that the regression errors are equal to the innovation errors.
   5. The new estimate for the effect of advertising is 0.5063. The constant is no longer estimated.
   6. The residuals are now stationary, meaning that our forecast intervals will likely be more accurate.
   7. The predicted sales are equal to 85.43
2. Piecewise linear regression with ARMA errors
   1. The model has ARIMA (2,0,0) errors, a negative downward trend until the knot point and upward slope after that.
   2. The forecasts initially decrease and then slowly increase.
3. Motel dataset
   1. Done in R
   2. Done in R. The CPI is included in the data.
   3. Logarithms should be taken due to the fact that the seasonal variation is not uniform throughout the series.
   4. The model chosen is ARIMA (0,1,1) (0,1,2). The ARIMA model for the errors also contains seasonal differencing due to the nature of the data. Information contained in previous seasons is useful to forecasting more recent errors.
   5. Done in R
4. Gasoline with ARMA errors
   1. After decomposing the data and extracting the trend-cycle component, it appears that there is a structural change in the series between 2005 and 2010, another one between 2010 and 2015, and one at the end of the series. The best model had K = 12. For time sake, I assumed that the nodes were:
      1. 2007.5
      2. 2012.5
      3. 2016
   2. Auto-arima takes forever
5. Electricity consumption.
   1. The innovation errors are modeled as an ARIMA (1,1,0) (0,1,2) [12]
   2. The positive coefficients mean that as the square root of monthly heating degrees increases by 1, we expect the log of monthly kilowatt hour to increase by 0.0077; if the square root of monthly cooling degrees increases by 1, we expect the log of monthly kilowatt hour to increase by 0.0208.
   3. Done in notebook.
   4. If we have temperature estimates for future months, we can replace the values to forecasts for electricity use.
   5. The problem is that the model without ARIMA errors will likely contain autocorrelated residuals. When regression models contain autocorrelated residuals, the estimated coefficients are biased because they are correlated with the error term; spurious regression might arise, etc. Taking account of possible residual autocorrelation alleviate or solve those problems.
6. Retail series analysis
   1. Using Auto.arima, the model with 6 Fourier terms had the smallest AICc. The model had ARIMA (1,0,4) (2,0,1) [12] errors
   2. The best model does not contain white-noise residuals
   3. Done in R.