

Homework 3

IE 529

Predictive Analytics

Due: June 05, 2020 (23:59)

This assignment is to be carried out individually. You can discuss your solutions with your friends, but please write your solutions individually. Submit your solution, including all the code scripts, as a single PDF or JPYNB file via LMS. (This exercises are from the textbook. All data are posted at LMS.)

Late Homework Policy: *If you submit your homework late up to 7 days, there will be a 20% penalty. Any homework will not be accepted after 7 days.*

1. The gasoline series consists of weekly data for supplies of US finished motor gasoline product, from 2 February 1991 to 20 January 2017. The units are in “million barrels per day”. Consider only the data to the end of 2004.
 - a. Fit a harmonic regression with trend to the data. Experiment with changing the number Fourier terms. Plot the observed gasoline and fitted values and comment on what you see.
 - b. Select the appropriate number of Fourier terms to include by minimizing the AIC value.
 - c. Check the residuals of the final model.
 - d. Forecast the next year of data.
 - e. Plot the forecasts along with the actual data for 2005. What do you find?
2. Consider monthly sales and advertising data for an automotive parts company (data set advert).
 - a. Plot the data.
 - b. Fit a standard regression model $y_t = a + bx_t + \varepsilon_t$ where y_t denotes sales and x_t denotes advertising using ordinary least squares.
 - c. Show that the residuals have significant autocorrelation.
 - d. What difference does it make you use the SARIMAX function instead where residuals are assumed to come from ARIMA (0,0,0)?
 - e. Refit the model using `auto_arima()`. How much difference does the error model make to the estimated parameters? What ARIMA model for the errors is selected?
 - f. Check the residuals of the fitted model.
 - g. Assuming the advertising budget for the next six months is exactly 10 units per month, produce and plot point sales forecasts for the next six months.

3. We fitted a harmonic regression model to part of the gasoline series in question 2. We will now revisit this model, and extend it to include more data and ARMA errors.
 - a. Using ordinary least squares, fit a harmonic regression with a piecewise linear time trend to the full gasoline series. Select the appropriate number of Fourier terms to include by minimizing the AIC value.
 - b. Now refit the model using `auto_arima()` to allow for correlated errors, keeping the same predictor variables as you used in part a.
 - c. Check the residuals of the final model using the `plot_diagnostics()` function. Do they look sufficiently like white noise to continue? If not, try modifying your model, or removing the first few years of data.
 - d. Once you have a model with white noise residuals, produce point forecasts for the next year.