CSC542/CSC642 Statistical Learning with Applications

Assignment #7

Melvin Maria

1. According to the author, smoothing splines are very commonly used. Based on what the paper says:
   a.) For what orders are they defined?

      - Normally smoothing splines are only defined for an odd polynomial order k.

   b.) Which is the most common value for k?

      - Most common: k=3

   c.) What type of penalty use smoothing splines? Have you seen this in another method in this class?

      - Smoothing splines uses a (squared) L2 norm penalty, which is equivalent to ridge regression method that we have seen in class.

   d.) In the empirical comparisons, which model performs best? What happens as we increase the degrees of freedom (df)?

      - From the hills example, the cubic trend filtering seems to provide a better estimate as it looked to adapt well to the levels of smoothness on left and right side of the domain. Looking at the df, increasing it may improve the fit, but it can overfit and under smooth other regions.

2. The author shows several methods in lasso form.

   a.) Which methods are these?

      - Locally adaptive regression spline in lasso form
      - Trend filtering in lasso form.

   b.) What penalty is used in this case?

      - These methods use the L1 norm penalty.

   c.) Which method performs better in the empirical comparisons in 3.4?

- For both methods the tuning parameters of a certain order, the estimators converge asymptotically at a fast rate, meaning they are too alike to determining which one performs the best.

d.) What is observed for small values of l?

- We start to notice the slightly differences on the two estimators.

3. The author examines astrophysics data.

a.) Briefly describe the data used and what the goal of the analysis is.

- The data used is from a quasar spectrum(n=1172), which are the one of the most luminous celestial bodies across the universe, goal was to estimate the underlying function.

b.) Briefly describe the experimental setting (i.e., methods used, parameters, frameworks/packages, tuning, comparison metric).

- For the estimation the methods used required 3 splines, trend filtering, smoothing (cubic order) and wavelet smoothing (4 vanishing moments) from the wavethresh package in R. Each method used over 146 values of degrees of freedom, to compare the three methods: the authors computed their average squared error loss to the true function, over 20 draws from the simulated model.

c.) What method performs best and why, according to the author?

- According to the author smoothing splines was outperformed by trend filtering for lower degrees of freedom.
- Wavelet smoothing was not competitive in terms of squared error loss, although in theory it achieved the same (minimax) rate of convergence as the trend filtering.
- In the end the split smoothing spline estimator was fitted over a total of 146 · 146 = 21,316 values of degrees of freedom which gave an advantage when compared to the other methods resulting in a lower squared error loss.