STAT 542/CS 598 Fall 2018

Coding Assignment 3

Due Thursday, November 15, 11:30 p.m. (PT)

1. (**Bonus Credit: 2pt**) Write your own functions to use LOO-CV and GCV to select the optimal span for `loess`. Check [Coding3_BonusCredit.html] on Piazza.

   - Test your code on data set `Coding3_Bonus_Data.csv`, which can be downloaded from Piazza.
   - Report your CV and GCV for 15 span values: 0.20, 0.25, . . . , 0.90.

   What you need to submit?

   A PDF file (maximum two-page) and the R/Python code that produces the PDF file. Your code will be run in a directory that has the file `Coding3_Bonus_Data.csv`.

   - The PDF file should contain the 15 values for LOO-CV and 15 values for GCV.
   - In [Coding3_BonusCredit.html], I include some figures, but you do not have to include them in your PDF file. The essential part of your PDF file are the LOO-CV and GCV values.
   - Students are allowed to use R/Python code to generate Markdown file in PDF. Since the file size is restricted to be two pages, suggest to hide your code and only display the results.

   - Name your R/Python file (Rmd files are allowed) starting with `Assignment_3_Bonus_xxxx_netID..`, where "xxxx" is the last 4-dig of your University ID.

     For example, the submission for Max Y. Chen with UID 672757127 and netID mychen12 would be named `Assignment_3_Bonus_7127_mychen12_MaxChen.R`. You can add whatever characters after your netID.

   - Name the PDF file (submitted along with your code) similarly, starting with `AssignmentOutput_3_Bonus_xxxx_netID...pdf`, where "xxxx" is the last 4-dig of your University ID. You do not need to worry about the name of the Markdown file generated by your code.

2. (**3pt**) Implement the Baum-Welch algorithm and the Viterbi algorithm for a hidden Markov model (HMM) with **Two** hidden states (A or B) and whose outcome is a discrete random variable taking **Three** unique values. Check [Coding3_HMM.html] on Piazza.

   - Test your code on data set `Coding3_HMM_Data.csv`, which can be downloaded from Piazza, with number of iterations for EM to be 100.
   - Report the output from your Baum-Welch algorithm, and save your output from your Viterbi algorithm.

   What you need to submit?

   A PDF file (maximum two-page) and the R/Python code that produces the PDF file. Your code will be run in a directory that has the file `Coding3_HMM_Data.csv`.

   - Report the estimated transition matrix $A_{2\times 2}$ and emission matrix $B_{2\times 3}$ from your Baum-Welch algorithm with 100 iterations in the PDF file.
   - Save the output, a sequence of length 500 taking values either "A" or "B", from your Viterbi algorithm in a file named `Coding3_HMM_Viterbi_Output.txt`. Then we will compare your result with the sequence from `Coding3_HMM_True_Viterbi_Output.txt`.
   - Students are allowed to use R/Python code to generate Markdown file in PDF. Since the file size is restricted to be two pages, suggest to hide your code and only display the results.

   - Name your R/Python file (Rmd files are allowed) starting with `Assignment_3_HMM_xxxx_netID..`, where "xxxx" is the last 4-dig of your University ID.

     For example, the submission for Max Y. Chen with UID 672757127 and netID mychen12 would be named `Assignment_3_HMM_7127_mychen12_MaxChen.R`. You can add whatever characters after your netID.

   - Name the PDF file (submitted along with your code) similarly, starting with `AssignmentOutput_3_HMM_xxxx_netID...pdf`, where "xxxx" is the last 4-dig of your University ID. You do not need to worry about the name of the Markdown file generated by your code.