ECES-434 Homework 4

Due Tuesday, November 19th, 2019

For this homework assignment, you will need to acquire a new python library, spectrum, that will allow you to compute LPC coefficients. With your anaconda terminal loaded, run the following commands:

- conda config –add channels conda-forge
- conda install spectrum

This assignment should be completed in the Homework4 Jupyter notebook which is available on Bb Learn. You will also need to download the voices folder to the same directory.

1. **Recognize speech or Wreck a nice beach!**

   Starting with the provided `lpcFrames` function, examine the 5 speech sound files provided on Bb Learn:

   `female1.wav` `male1.wav`

   `female2.wav` `male2.wav`

   `female3.wav`

   The files contain different people saying the phrase "Critical equipment needs proper maintenance". They are sampled at 16 kHz.

   (a) Early attempts at developing systems for automatic speech recognition (ASR) attempted to utilize formant frequencies to identify different phonemes. We can estimate the formant frequencies from the pole locations of the LPC polynomial $A(z)$ (the columns of the output from `lpcFrames`), using the functions `np.roots` and `np.angle`. Unfortunately, `np.roots` will only work on one column at a time, so write a python function to convert each column into its complex poles, and then find the angle of each pole. Of course, `np.angle` produces an angle in radians (where π is the Nyquist frequency), so also convert those to Hz and plot the pole angles vs. frame number on a single graph (per speaker).

   (b) Note any similarities or differences you see that may point to a method of performing ASR.

   (c) When using LPC for speech compression, quantization errors of the individual LPC coefficients α_k can lead to unstable filters, which is undesirable. Some smart engineers came up with a method where our LPC polynomial $A(z)$ of order P is converted to two other polynomials:

   $$P(z) = A(z) + z^{-(P+1)}A(z^{-1})$$
   $$Q(z) = A(z) - z^{-(P+1)}A(z^{-1})$$

   You can easily verify that

   $$A(z) = \frac{P(z) + Q(z)}{2}.$$

   What's not as obvious is that the roots of $P(z)$ and $Q(z)$ all lie on the unit circle, so they can be completely described by just their angles. Additionally, since the roots occur in complex conjugate pairs, only half of the frequencies of $P(z)$ and $Q(z)$ are needed. These frequencies are called *line spectrum frequencies* (LSFs). Fortunately, the spectrum library has a function that will do the calculation of LSFs for you, `poly2lsf`. Again, this function will only operate on a single column of data at a time, so you'll need to create a function that converts each column into the equivalent LSFs. Plot the LSFs vs. frame number for each speaker.

   (d) Do the LSFs appear better suited for ASR than the pole angles from part (a)?