EE679: Computing Assignment 2

Part 1: Linear Predictive Analysis

Due: Oct 24, 2019

The submission must include: method, important code fragments, labeled plots, discussion. Note: ACF computation (from waveform samples) must be written from scratch. Also, compute the real cepstrum using basic functions such as fft and log. You may use library functions if you wish for other blocks.

A. Synthesized vowel: Take (from your previous assignment) the synthesized vowel /a/ (formants: 730, 1090, 2440 Hz; bandwidths: 100 Hz) at two fundamental frequencies: 120 Hz, 220 Hz. Sampling rate = 8 kHz.

Take a 30 ms Hamming window, implement LP analysis on this single segment using LP orders 2, 4, 6, 8, 10 via the Levinson-Durbin recursion. Compute the gain, and plot the LP spectrum magnitude (i.e. the dB magnitude frequency response of the estimated all-pole filter) for each order "p". Show your true underlying 6-pole spectral envelope with the discrete harmonic components shown with vertical lines (emulating the ideal spectrum of a periodic signal). Compare this with the obtained plots to comment on the characteristics of the spectral approximations of different orders for the different F0 vowels.

B. Natural speech: Consider the speech signal in "machali.wav" (male voice), sampled at 8 kHz. Consider the following signal segments in the final word "pani": (1) /a/ (first half); (2) /n/; and (3) /s/ in the word "uska".

Use PRAAT to extract the above segments to separate .wav files for further analyses as below. (Note: for /s/, 16 kHz sampled audio is better.)

1. Compute and plot the narrowband spectrum of each sound using a Hamming window of duration = 30 ms (after pre-emphasis for the voiced sounds).

2. Using a 30 ms Hamming window centered in the segment of the waveform (use pre-emphasis only for the voiced sounds), implement the below for each of the 3 sounds:

(a) Compute the autocorrelation coefficients required for LPC calculation at various p = 4,6,8,10,12,20. Use the Levinson-Durbin recusion to compute the LP coefficients from the autocorrelation coefficients. Show the pole-zero plots of the estimated all-pole filter for p=6,10.

(b) Compute the gain and plot the LPC spectrum magnitude (i.e. the dB magnitude frequency response of the estimated all-pole filter) for each order "p". Superimpose each plot on the narrowband dB magnitude spectrum of part 1. Comment on the characteristics of the spectral envelope estimates.

(c) Plot error signal energy (i.e. square of gain) vs p.

3. Based on the 10th-order LP coefficients, carry out the inverse filtering of the /a/ vowel segment and of the unvoiced sound /s/. Obtain the residual error signal in each case. Can you measure the pitch period of the voiced sound from the residual waveform? Use the acf to detect the pitch. Compare the acf plots of the original and residual signals. Plot the magnitude spectrum of each of the residual signals, and comment on it.

4. LP re-synthesis: We analysed the natural speech sounds /a/, /n/, /s/ in the earlier questions. Using a suitable set of parameter estimates for each sound as obtained there, we wish to reconstruct each of the sounds.

   That is, use the best estimated LP filter with an ideal impulse train of the estimated pitch period as source excitation (for the voiced sounds). Carry out de-emphasis on the output waveform. For the unvoiced sound, use a white noise signal as source excitation. Set the duration of the synthesized sound to be 300 ms at 8 kHz sampling frequency (use 16 kHz for the /s/), and view the waveform as well as listen to your created sound.

   Comment on the similarity with the original sound. What would be a good application for this analysis-and-synthesis system, and how exactly would it help?

Part 2: Cepstral Analysis

Due: Oct 31, 2019

1. Next, consider the synthetic signal (for /a/) reconstructed from LP coefficients (p=10) and pulse train in Part 1 B: Q4. Compute the real cepstrum and obtain the spectral envelope (dB) via cepstral liftering. Compare this estimated spectral envelope with the true LP magnitude spectrum. Observe and comment on the changes in cepstral estimate of the spectral envelope with different duration lifters.

2. Finally, we wish to carry out the cepstral analyses of the same natural speech sounds. Obtain the real cepstrum from a 30 ms segment for each of the phones (of the natural speech as in Part 1 B). Estimate the voicing and pitch of the segment from the real cepstrum. Use cepstral liftering to obtain the spectral envelope (dB) in each case.

   Compare it with the corresponding LP (p=10) magnitude spectrum obtained previously (part B.2) by superposing both on the actual magnitude spectrum of the windowed signal. (Obtain the vocal tract magnitude response of /s/ sampled at 16 kHz using LP order = 18.)