CS571 Project Report

VIKAS SINGH,ROLL NUM-t21017

23/11/2021

1 Summary

Linear prediction(LP) is one of the most important tools in speech analysis .The philosophy behind linear prediction is that a speech sample can be approximated as a linear combination of past samples .Then by minimizing the sum of the squared difference between the actual speech samples and the linearly predicted ones over a finite interval, a unique set of predictor coefficient can be determined .LP analysis decomposes the speech two highly independent components,the vocal tract parameters(LP coefficients)and the global excitation(LP residual) .it is assumed that speech is produced by exciting a linear time-varying filter(the vocal tract) by random noise for unvoiced speech segments ,or a train of pulses for voiced speech . The most general predictor form in linear prediction is the autoregressive moving average model where the speech signal s(n) is modelled as a linear combination of the past outputs and the present and past inputs .It can be written mathematically as

$$s(n) = -\sum_{K=1}^{P} a_k s(n-k) + G\sum_{l=1}^{q} b_l u(n-l).$$

where, gain G is parameter of the filter Equivalently in frequency domain the transfer function of the linear prediction speech model is

$$H(z) = (1 + \sum_{l=1}^{q} b_l z^{-} l)/(1 + \sum_{K=1}^{P} a_k z^{-} k).$$

H(z) is referred to as a pole-zero model .The zeros represents the nasals and poles represent the resonance(formats) of the vocal-tract. when

$$b_l = 0$$

H(z) becomes all pole or autoregressive model.

$$H(z) = 1/(1 + \sum_{K=1}^{P} a_k z^{-} k)$$

generally all pole-pole model is preferred for most application .The zeros are approximated modelled by including more poles and the location of poles considerably more important perceptually than the location of zero ,moreover it is easy to solve all pole model.

The error signal or the residual signal e(n) is the difference between the input speech and the estimated speech.

$$e(n) = s(n) + \sum_{K=1}^{P} a_k s(n-k).$$

There are two widely used methods for estimating the LP coefficients(LPCs) (1). Autocorrelation (2). Covariance in both the methods choose the short term filter coefficients

$$a_k$$

in such a way that energy in the error signal can be minimized . FOR the speech processing tasks, the Autocorrelation method is used because its computational efficiency and inherent stability are good.

2 Introduction

Linear predictive coding(LPC) is a method used mostly in Audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form ,using the information of a linear predictive model. LPC is most widely used method in speech coding and speech synthesis . It is a powerful speech analysis technique , and useful method for encoding good quality speech at low bit rate.

3 Solution

(1). Import various useful packages
(2). Input speech ,fs=soundfile.read(inputAudio.wav)
(3). Window=scipy.windows.hamming(fs,False)
(4). function frameblocks(signal,window,o=0.5) create input speech into frames returns frames
(5). For each frame:
dftframe=numpy.fft.fft(frame)
plot numpy.log10(abs(dftframe))
lpccoeff=librosa(frame,order),returns order+1 length vector
w,h=scipy.freqz(lpccoeff,1)
plot(w,numpy.log10(abs(h))
poles=1
residual=signal.Ifilter(coefficients,poles,frame) inverse filter
synthetic=signal.Ifilter([1],coefficients,residual)
return all synthetic frames
(6). functionaddframeblocks(signal,windows,o=0.5) , returns the signal ,return output signal
(7). soundFile.write(outputsignal,fs)
(8). end

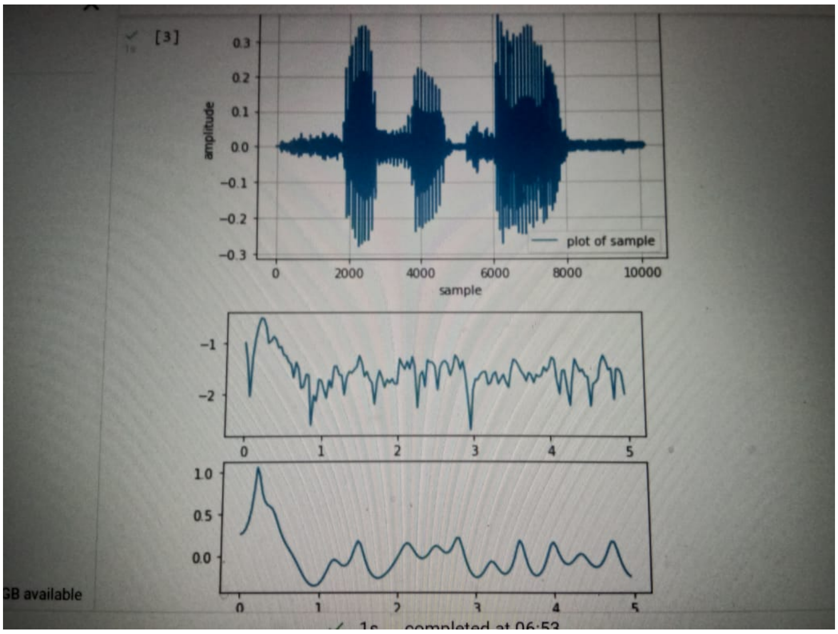
3.1 Assumptions

(1). it is assumed that speech is produced by exciting a linear time-varying filter(the vocal tract) by random noise for unvoiced speech segments ,or a train of pulses for voiced speech
(2). Order of linear filter is greater than 1 and integer as well.
(3). librosa.lpc(y,order),where y should not be ill-conditioned

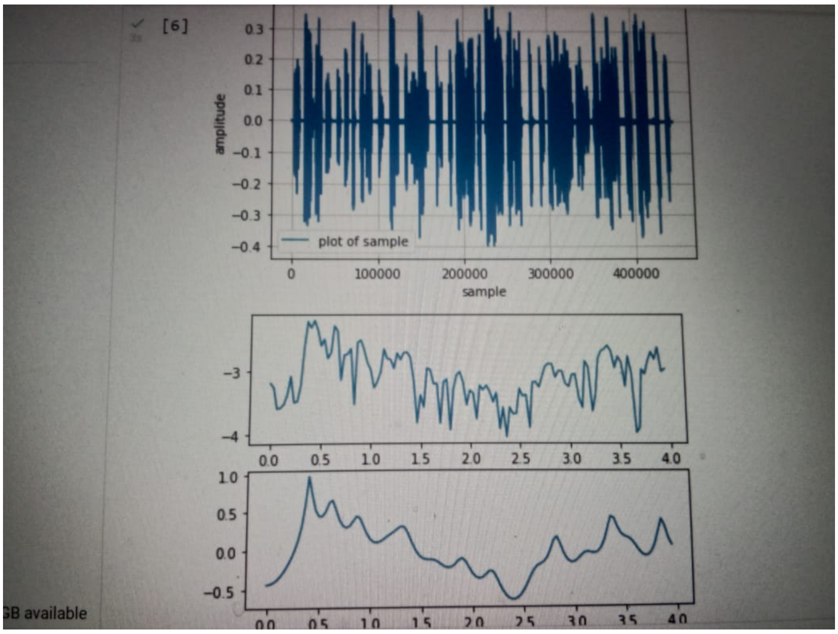
3.2 Algorithms used

1. Divide the speech signals into hamming windowed frames. 2. The following operations are to done framewise. 3. Perform LPC analysis. plot the DFT spectrum and LPC spectrum on the same plot. 4. obtain the inverse filter. 5. Using inverse filter, obtain the LPC residual signal. 6. Use the LPC residual to excite the LPC filter and produce a speech frame. 7. perform overlap-add to obtain back the speech signal. play back the speech signal .

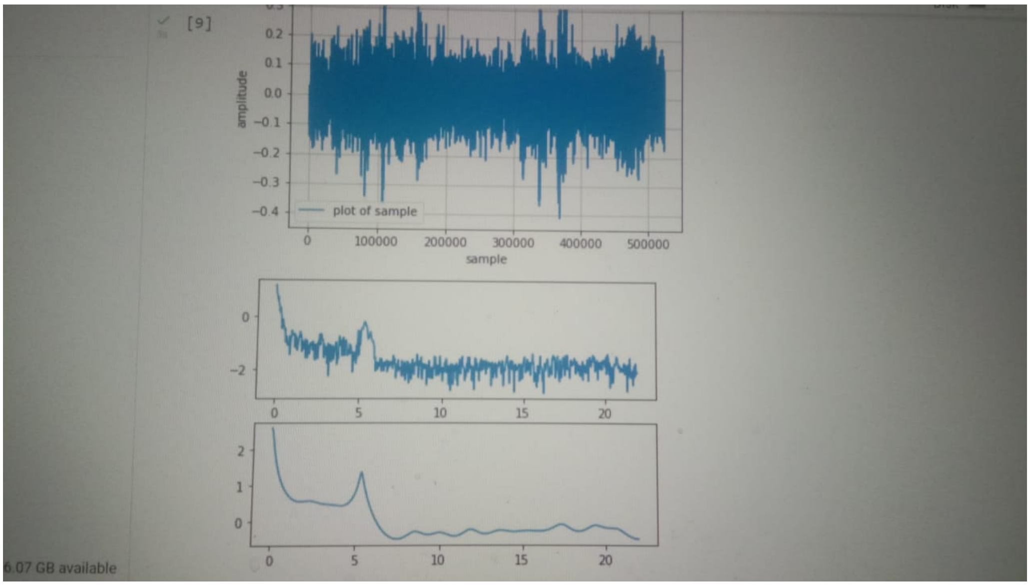
4 Results and analysis



should.output



phone.output



amador.output

5 Conclusion

I am getting DFT of given .wav file and LPC envelop which is peak detection.i am using here librosa.lpc(y,order),this function applies Burgs method to estimate coefficients of linear filter on y of order.
raises : (1)parameter error-
if y is not valid audio as per librosa.util.validaudio
if order¡1 or not integer
(2). Floating point error - if y is ill condition.

6 Project Github page

https://github.com/vikassssssssss/Linear-predictive-coding-of-speech.git

7 Bibliography

Theory and Applications of Digital Speech Processing by Rabiner and Schafer,
Prentice Hall Rabiner's Speech processing website: https://web.ece.ucsb.edu/Faculty/Rabiner/ece259