MATLAB Exercise – LPC Error Spectra

Program Directory: `matlab_gui\lpc_error`
Program Name: `lpc_error_GUI25.m`
GUI data file: `lpc_error.mat`
Callbacks file: `Callbacks_lpc_error_GUI25.m`
TADSP: Section 9.6, pp. 527-531, Problem 9.34

This MATLAB exercise shows the (flat spectrum) nature of the LPC error signal for a typical speech frame (voiced speech).

LPC Error Spectra – Theory of Operation

This exercise processes a user-designated frame of speech using either the autocorrelation method or the covariance method of LPC analysis. Using the resulting LPC model, the exercise computes the error signal by filtering the speech input by the inverse filter.

LPC Error Spectra – GUI Design

The GUI for this exercise consists of two panels, 4 graphics panels, 1 title box and 13 buttons. The functionality of the two panels is:

1. one panel for the graphics display,
2. one panel for parameters related to the linear prediction analysis, and for running the program.

The set of four graphics panels is used to display the following:

1. the window-weighted frame of speech used for LPC analysis,
2. the LPC error signal,
3. the log magnitude spectrum from the STFT of the frame of speech, along with the LPC log magnitude spectrum obtained from LPC analysis of the same frame of speech,
4. the log magnitude spectrum of the LPC error signal.

The title box displays the information about the selected file along with the set of LPC analysis parameters. The functionality of the 13 buttons is:

1. a pushbutton to select the directory with the speech file that is to be analyzed using short-time analysis methods; the default directory is 'speech_files',
2. a popupmenu button that allows the user to select the speech file for analysis,
3. a pushbutton to play the current speech file,
4. a popupmenu button that allows the user to select either the Autocorrelation method or the Covariance method for LPC analysis; (the default is the Autocorrelation analysis method),
5. a popupmenu button that lets the user choose either a Hamming or Rectangular window as the short-time LPC analysis window for the autocorrelation method of LPC analysis; (default is the Hamming window),
6. a popupmenu button that allows the user to select the window for STFT of the signal (either Hamming or Rectangular windows can be chosen; default is the Hamming window),
7. an editable button that specifies the frame duration, L_m, (in msec) for short-time analysis; (the default value is $L_m = 40$ msec),

8. an editable button that specifies the frame shift, R_m, (in msec) for short-time analysis; (the default value is $R_m = 10$ msec),
9. an editable button that specifies the LPC system order, p; (the default value is $p = 16$),
10. a pushbutton to choose the single frame starting sample, `ss`, using the iterative method described in the section below; this starting sample defines the current frame of the speech signal,
11. a pushbutton to run the analysis code and display the signal processing results using the current frame of the speech signal; this button can be pressed and used as often as desired, changing one or more analysis parameters while keeping the frame starting sample the same,
12. a pushbutton to run the analysis code and display the signal processing results using the next frame of signal; i.e., the frame with starting sample set to `ss+R` where `R` is the frame shift in samples; this button can be pushed repeatedly to provide a frame-by-frame analysis,
13. a pushbutton to close the GUI.

Interactive Method of Defining the Speech Analysis Frame Starting Sample

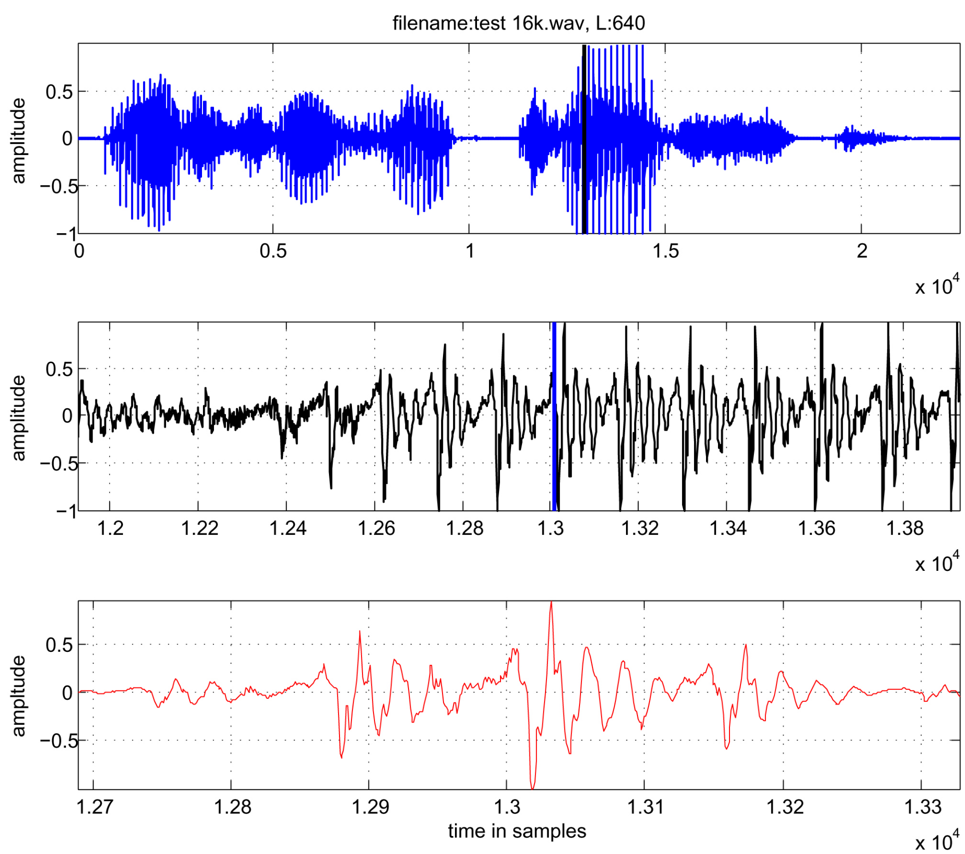

Figure 1: Sequence of waveform plots defining how the user can interactively choose a starting sample for the current analysis frame.

Several MATLAB Exercises rely on frame-based analysis methods where the user needs to specify both the speech file for analysis, and the starting sample of the speech analysis frame of interest. The method that we have chosen to define the frame starting sample is an interactive analysis which homes in on an appropriate analysis frame in a series of steps. The operations of this interactive method for determining the starting sample of the speech analysis frame for autocorrelation analysis proceed as follows:

1. In a specified graphics frame (or figure sub-frame) a single line plot of the entire speech waveform is obtained, as illustrated at the top panel of Figure 1. A graphics curser then appears allowing the user to move the cursor to the region of speech that is of interest for specifying the current analysis frame. A solid vertical cursor is shown at the place selected by the user. For the example of Figure 1 the cursor location is approximately sample 13000, as indicated by the solid red bar.

2. In another specified graphics frame (or figure sub-frame) a plot of the speech signal over a region that is about ± 1000 samples around the location of the cursor in the previous step; i.e., from sample 12000 to sample 14000. A second graphics cursor appears allowing the user to move the cursor to the exact starting sample of interest (to within the resolution of the display) for specifying the current analysis frame, as illustrated in the middle graphics panel of Figure 1. Here the cursor is again shown in the area of sample 13000.

3. The current analysis frame is then defined as the frame of speech from the starting sample of step 2 minus half the window length, to the starting sample of step 2 plus half the window length. The designated analysis frame is then weighted by the analysis window (Hamming in the case here) and plotted in the bottom graphics panel.

It should be clear that the three steps of the above process for choosing an analysis frame can be implemented in either a single graphics panel or frame (by simply overwriting the graphics panel with the new speech signal) or in a series of graphics panels or frames. The current exercise uses one of the 8 graphics panels and overwrites the speech waveform plot at each step of the analysis. This process is a very useful and efficient one for choosing a region of interest within the speech signal, and then homing into a particular analysis frame using the steps outlined above.

LPC Error Spectra – Scripted Run

A scripted run of the program 'lpc_error_GUI25.m' is as follows:

1. run the program 'lpc_error_GUI25.m' from the directory 'matlab_gui\lpc_error_spectra',

2. hit the pushbutton 'Directory'; this will initiate a system call to locate and display the filesystem for the directory 'speech_files',

3. using the popupmenu button, select the speech file for short-time feature analysis; choose the file 'test_16k.wav' for this example,

4. hit the pushbutton 'Play Speech File' to play the current speech file,

5. using the popupmenu button choose Autocorrelation as the method for LPC analysis,

6. using the popupmenu button choose Hamming for the STFT analysis window,

7. using the popupmenu button choose Hamming for the LPC frame analysis window (when using the autocorrelation method of LPC analysis),

8. using the editable buttons, choose initial values for the analysis parameters including 40 msec for frame length, L_m, 10 msec for frame shift, R_m, and 16 for LPC system order, p,

9. hit the 'Get Frame Starting Sample' button to interactively choose the initial analysis frame starting sample, `ss`, using the iterative method described in the section above; try to choose the starting sample as close to the value of 13190 so as to match the plotted results for this example exercise,

10. hit the 'Run Current Frame' button to initiate single frame analysis of the speech beginning at the current frame starting sample, `ss`; the results of LPC error analysis are shown in the various graphical plots; the 'Run Current Frame' button can be hit repeatedly after making changes in the analysis frame parameters,

11. hit the 'Run Next Frame' button to initiate single frame analysis on the next frame of speech, i.e., where the starting sample of the next frame is set to `ss+R`, where `R` is the frame shift in samples,

12. experiment with different choices of speech file, and with different values for L_m, R_m, p, STFT window type, LPC analysis window type (used for autocorrelation method of LPC analysis) and LPC analysis method,

13. hit the 'Close GUI' button to terminate the run.

An example of the graphical output obtained from this exercise using the speech file 'test_16k.wav' is shown in Figure 2. The graphics panels show the speech frame used for STFT and LPC spectrum analysis (top graphics panel), the LPC error signal (second graphics panel), the log magnitude responses of the STFT and the LPC system (third graphics panel), and the log magnitude response of the LPC error signal (bottom graphics panel). The flatness of the log magnitude spectrum of the error signal is clearly seen in Figure 2.

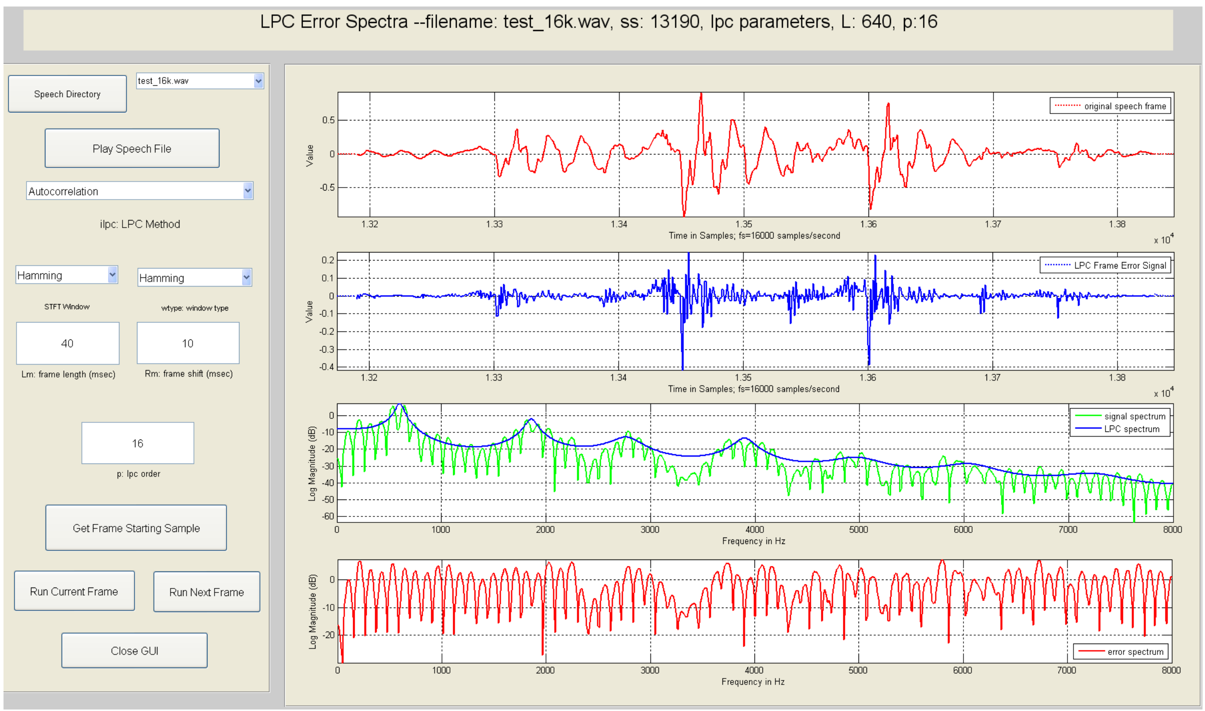

Figure 2: Example of LPC error analysis using a voiced frame of speech analyzed by both the STFT and the LPC autocorrelation method. The upper graphics panel shows the window-weighted speech frame; the second graphics panel shows the LPC error signal; the third graphics panels shows the log magnitude spectrums of the STFT and the LPC system; and the bottom graphics panel shows the log magnitude spectrum of the LPC error signal.

LPC Error Spectra – Issues for Experimentation

1. run the scripted exercise, above, using the speech file 'test_16k.wav', with signal processing parameters of frame starting sample, ss = 13103 (or a close value); autocorrelation method of LPC analysis, Hamming window with frame length of $L_m = 40$ msec, and LPC system order of $p = 16$. Using this section of speech, answer the following questions:

   - how closely does the error signal (shown plotted in the second graphics panel from the top) match the ideal excitation signal for a Hamming window weighted voiced section of speech?

- what is the excitation pitch period (in samples and in msec) for this section of speech?
- how closely does the error spectrum (shown plotted in the bottom graphics panel) match the ideal (flat spectrum) for this section of voiced speech?
- how would your answers to the previous questions change if we made the following changes to the signal processing parameters:
  - use the covariance method of LPC analysis instead of the autocorrelation method?
  - change the frame duration from $L_m = 40$ msec to $L_m = 80$ msec (i.e., analysis window of twice the standard length), or from $L_m = 40$ msec to $L_m = 20$ msec (i.e., analysis window of half the standard length),
  - change the LPC system order from $p = 16$ to $p = 32$ or to $p = 8$.

2. repeat the signal processing analysis on an unvoiced region of speech using the frame of the file 'test_16k.wav' beginning at ss=3195 (or reasonably close to this frame starting sample). Answer all the above questions for the unvoiced section of speech.

3. repeat the analysis of this MATLAB exercise using the synthetic speech file 'vowels_100Hz_edited.wav'. This file contains the set of 10 American English vowels synthesized using an ideal vocal tract impulse response excited by a periodic impulse train of period 100 samples. Use a frame of signal beginning at sample 200 for this exercise. Using the default signal processing parameters, examine the resulting error signal and error spectrum for this frame of speech. How much closer to the ideal excitation signals are the error signals from this analysis? What signal processing phenomenon explains the error signal for frame samples $m = 0, 1, ..., p - 1$ and for $m = L, L+1, L+2, ..., L+p-1$? What is the difference in the error signal between using a Hamming window and a rectangular window?

- using the speech file '1A.waV' select a starting sample in the vicinity of sample 5000. Run the LPC analysis using both the autocorrelation and the covariance method. Compare the LPC spectrums for frequencies above 4000 Hz. Why are the spectral matches so different in this region of the spectrum? (Think about the impact of using a Hamming or Rectangular window on the STFT analysis in a region of spectrum where there is a very low signal level.)

- using the speech file 's2.wav' set the frame length to $L_m = 30$ msec. Run the LPC error analysis using the autocorrelation method of LPC. Select a frame starting sample around sample 1600 (the region of the /IY/ vowel for the word 'thieves' in the utterance "Thieves who rob friends deserve jail".

  - observe the $p = 16$ samples at the beginning and end of the prediction error signal. Why is the prediction error large in these intervals? Why does the prediction error contain "impulses"? What is the spacing between impulses? How is the spacing between impulses related to the pitch period of the waveform, as seen in the top graphics plot?
  - compare the STFT of the signal to the STFT of the prediction error. How do they differ? In both cases you should see the same shape repeated periodically in frequency. What is the period, and what is the shape that is repeated?
  - now switch to the covariance method of LPC analysis, using the same speech frame as above. Why is the prediction error not large at the beginning and end of the window? Is the prediction error more or less impulsive than that obtained from the LPC autocorrelation analysis?