

Spectral Estimation

Please load J-DSP by pressing the  button of the NEEDS award J-DSP web page.

Problem 7-1: Classical Spectral Estimation

There are two types of classical spectral estimators: Correlograms and Periodograms. In this lab you will be using a **Sig Gen(L)**, the Long Signal Generator that provides some elementary framing capabilities. The framing schemes determine the flow of data into the rest of the system. The data is always generated in a frame-by-frame mode. You can choose the frame size and the degree of overlap. You can process and view frames *manually* i.e. each frame is generated one-at-a-time after you press the button >>. You can also process the frames *automatically*. In this case, some delay can be induced if desired by using the *Pause* button.

The spectral analyzers are stand-alone blocks and have their own graph and parameter-entry capabilities. In this exercise you will test three different methods, i.e., two classical and one parametric. We start first with the correlogram.

7-1 (a) Correlogram

Step – 1

Establish the **Sig.Gen(L)** block and generate colored noise by specifying *colored* in the input drop-down menu. Enter the following transfer function

$$H(z) = \frac{1}{1 + 0.81z^{-2}}$$

Note: 8192 samples of color noise data are automatically generated.

Step – 2

Connect the output of the **Sig.Gen(L)** (long signal generator) to the **autocorr** block. Change the *lag* field to 16 in order to compute 16 lags. Connect **lag window** to this block. This should be connected to the **SymCorr** to make the correlations symmetric and hence produce a real-valued correlogram. Connect this block to the **Correlogram** block. The graph panel is in-built in the **Correlogram** dialogue box.

Problems

7.1 Save the block diagram as *graph1.gif*

7.2 Compute the autocorrelations for the *current frame* and for the *entire signal* (up to the current frame). Notice the changes in the correlogram plot for both the cases. Which of the above two cases provide better estimates? Comment on the results you obtained. Save the correlogram plots of a specific frame (lets say frame 15) for the above two cases as *graph2.gif* and *graph3.gif* respectively.

7.3 Change number of *lags* in the autocorrelation block to **4** and then **64**. Notice the variations in the resolution of the correlation estimates. Save the correlogram plots for the above two cases (lags 4 and 64) as *graph4.gif* and *graph5.gif*

7.4 Select *hamming* in the lag window and observe the changes in the estimated PSD. Comment on the changes you noticed. Save the correlogram plots obtained with and without the hamming window, as *graph6.gif* and *graph7.gif*

Note

The frequency response of the above filter can be viewed independently by selecting the **filter** block, add **coefficient** block to the lower input and connect the **filter** block's top output to the **Freq Resp** block. Note that the ideal spectral density has the same shape as the frequency response.

$$|H(e^{jq})|^2 = \left| \frac{1}{1 + 0.81e^{-2jq}} \right|^2 \quad \text{ideal PSD}$$

7-1(b) Periodogram

Step –1

Follow the step – 1 of 7-1(a), and connect the output of the **Sig.Gen(L)** to the **Periodogram** block. The graph panel is in-built in the **periodogram** dialogue box.

Problems

7.5 Change the *overlap* from 0% to 50% in the **Sig.Gen(L)** block. Click on the '>>|' button (*No Pause*) frame-by-frame processing. Change the *periodogram* from *sample spectrum* to *Welch* and

observe the changes. Save the periodograms obtained for sample spectrum case and Welch case as *graph8.gif* and *graph9.gif*

7.6 Change the window in the **periodogram** block to **Hamming** window. Click on the '>>' button

(*Pause every frame*) and observe the results. Save the periodogram plots (for frame 15) with and without hamming window as *graph10.gif* and *graph11.gif* respectively.

7.7 Select the **smoothing** property (Daniell periodogram). Set the smoothing parameter to **0**. Click on the '>>' button (*Pause every frame*) and observe the results. Comment on your results. Save the periodogram plot obtained as *graph12.gif*

Problem 7-2: AR Spectral Estimation

Step – 1

Follow the step – 1 of 7-1(a), and connect the output of the **Sig.Gen(L)** to the **AR Est.** block. The graph panel is in-built with in the **AR Est.** dialogue box.

Step – 2

The **order** of the **AR Est.** is set to be **10**. Set **normalization** to unbiased.

7.8 Set the **order** of the **AR Est.** to **2**, **16** and notice the resolution of the AR estimates obtained. Comment on your results. Save the graphs obtained for orders **10**, **2** and **16** as *graph13.gif*, *graph14.gif* and *graph15.gif* respectively

7.9 Change the coloring filter transfer function in Sig.Gen(L) to

$$H(z) = \frac{1}{1 + 0.25z^{-2}}$$

and observe the changes in the AR Estimates obtained for orders 10, 2 and 16 respectively. Save the graphs as *graph16.gif*, *graph17.gif* and *graph18.gif* respectively. Comment on your results.

Problem -7 (Linear predictive coding¹)

For this problem, a separate report should be typed. Label the figures and tables sequentially and aptly. Consider,

- Input: *Male speaker-1*
- Frame size: 256 (Total number of frames = 64, i.e., 16384 samples)
- Overlap: 0%

Use the *Sig. Gen(L)* block for the speech input.

PART-A (FFT spectra of speech frames)

[7-1] Use the *FFT* block to compute the FFT spectra of speech. Observe the spectra for all the 256-point frames of speech and present in your report at least one plot of time and one corresponding plot of frequency-domain data for a voiced, unvoiced and a mixed segment (total 6 plots). Use *Plot2* block to plot the time-domain and frequency-domain plots together.

PART-B (Visually recognize the pitch period)

In this problem, you will measure the pitch period and the voicing of the speech frames. The period of a strongly voiced signal is associated in a reciprocal manner to the fundamental frequency of the corresponding harmonic spectrum. So if the pitch period is T , the fundamental frequency is $1/T$. Measure the pitch period visually from the time-domain plot in terms of the number of samples in a pitch period. If the segment is unvoiced/pause, enter infinity for the pitch period and hence zero for the fundamental frequency. If the segment is mixed, do your best to obtain an estimate of the pitch period, or set the pitch to infinity. For the voicing decisions do the following: when the segment is voiced enter '1' in the second column. If speech pause (no speech present) enter '0', if unvoiced enter '0.25' and if mixed enter '0.5'.

¹ This problem was given as a class project in EEE 407 Digital Signal Processing class, Fall-2002.

[7-2] For the first one second of speech (roughly 30 frames), tabulate the following: This problem is done by visual inspection of the frames. Use the zoom feature in *Plot* block to visualize the signal better. Use MS-EXCEL to create the table as that will facilitate plotting some of the required figures that you need to provide in this problem.

Table-3: Pitch period and voicing for the first one second of speech

Speech frame number	Voiced/unvoiced/mixed/pause	Pitch (No. of samples)	Energy of frame	Fundamental frequency(Hz)
1				
2				
3				
:				
:				
31				

[7-3] From the table above, plot the *fundamental frequency* as a function of the frame number for all thirty one frames. This is called the pitch frequency contour.

[7-4] From the table above, plot the *voicing* as a function of the frame number for all the thirty one frames.

[7-5] From the table above, plot the frame energy in dB as a function of the frame number for all thirty one frames.

[7-6] Describe in a page what have you learned with regard to the pitch period, the local energy of speech, and the voicing. Justify with your results.

PART-C (The LPC and the FFT spectra)

Use the *Sig. Gen(L)*, *LPC+*, and *FFT* blocks. Set up the block diagram as shown in Figure D.1.

Use the *Freq. Resp.* and *Plot* blocks to view the LPC spectra and the FFT spectra respectively.

[7-7] Use the frame-by-frame processing feature of the *Sig.Gen(L)* block to compute the LPC spectra of all the frames. From the LPC spectra, measure the frequencies of the first three formants, F_1 , F_2 , and F_3 . Use MS-EXCEL to create the table as that will facilitate plotting some of the required figures that you need to provide in this problem. Tabulate as follows (Give the frequencies in Hz):

Table-4: Formant frequencies for the first one second of speech

Speech frame number	Speech/pause	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
1				
2				
3				
:				
:				
31				

Note: The input speech is sampled at 8 KHz. If there are pauses, then just include whatever the frequencies of the peaks are. In column-2, enter '1' for speech and '0' for pauses.

[7-8] Plot the three formants across the frame number. These will be the formant contours. Use different line types or colors to discriminate the three contours.

[7-9] Use the *Plot* block to plot the FFT spectra, and the *Freq. Resp.* block to plot the LPC spectra. Note that the LPC spectra look like a smoothed version of the FFT spectra. We

call the spectrum from the LPC, the spectral envelope. Describe in a page what have you learned with regard to spectra of speech and the all-pole filter. Justify your answer.

PART-D (Speech synthesis using the LPC)

In this problem, we calculate the prediction residual and will use it to excite an all-pole filter that will synthesize the speech signal. The synthesis is done frame-by-frame basis, and all frames will be concatenated to produce a new speech record which we can then listen to and evaluate. Use *Sig. Gen(L)*, *LPC+*, *filter*, and *FFT* blocks in J-DSP. Use the *Snd. Player* block to view the synthesized (and concatenated) speech.

[7-10] Plot the excitation vector (residual) and its corresponding dB spectrum for a voiced and an unvoiced frame (total of 4 plots).

[7-11] Provide plots of the entire original and reconstructed speech for the speech file.

[7-12] Compute the segmental signal-to-noise ratio (SNR) in dB for the 31 frames. Listen to the reconstructed signal and provide a subjective score on a scale of 1- 5.

[7-13] Describe in a page what have you learned with regard to spectra of residual, quality of reconstruction.

PART-E (Speech synthesis using the LPC)

In this problem, we will down-sample and up-sample the prediction residual by different factors and see the effects on speech synthesis. The reconstructed signal is assessed in terms of an SNR and the subjective evaluation. The subjective scores are on a scale of 1 to 5 – one is bad and five is excellent.

Note: Cellular telephone quality is usually rated 3.5, wire-line telephone quality is rated at 4.2, and, voice in military radios is rated at 2.8.

[7-14] For every frame down-sample the prediction residual by a factor of 4. Then up-sample it by a factor of 4. Use the *downsamp* and the *upsamp* blocks to perform down-sampling and up-sampling respectively. The down-sampling anti-aliasing filter and the up-sampling reconstruction filter are not used. In other words, the processed signal, $e(n)$, in

every frame after the rate changes will now contain several zero samples (3 out of 4 samples will be zero). Therefore we will have only 64 non-zero samples in each 256-point residual frame. Hence, this signal contains less information than the original prediction residual. Compute the segmental SNR in dB and provide a subjective score on a scale of 1 to 5.

[7-15] Repeat the problem [7-15] for the down-sample and up-sample factor of 16.

[7-16] In this problem, you will design a Butterworth filter to realize the anti-aliasing and the reconstruction filters. Use *FIR Design* in J-DSP to design the Butterworth filter. Choose your own specifications. Give the specifications of the filter in your report and the transfer function of the filter used. Also, give the plots of its frequency response in dB. Use these filters and repeat problem [7-16]. Compute the SNR in dB and provide a subjective score on a scale of 1 to 5. Tabulate the above details as follows:

Table-5: The SNR and the subjective score results

Problem	SNR (dB)	Subjective score (1 to 5)
[7-12]		
[7-14]		
[7-15]		
[7-16]		

[7-17] Describe in a page what have you learned with regard to reconstruction with down-sampled/up-sampled residual spectra, quality of reconstructed signal, and effect of anti-aliasing and reconstruction filters. Justify.