Signals and Systems

Laboratory Number 3: Basic System Concepts and Speech Processing

Format for Lab Report

1. Title Page
2. Abstract - maximum of 300 words
3. Objectives
4. Introduction and Motivation
5. Background Material (includes any theoretical background)
6. Results and Discussion
7. Conclusion
8. Appendices - put listings of MATLAB code in the Appendices. Put any other Appendices you see fit.

Questions 1 and 2 deal with MATLAB functions that are part of the symbolic math toolbox.

Question 1 Transfer Function and Impulse Response

Consider the transfer function of an LTI system given by:

$$H(s) = \frac{s^2 + 5s + 6}{s(s+1)(s^2 + 10s + 50)}$$

(a) Use MATLAB to determine the zeros and poles of H(s) and plot them in the s-plane. Use a circle to denote a zero and an x to denote a pole. You will find the ‘roots’ function helpful.

(b) Plot the impulse response h(t) of the system. The ‘impulse’, ‘conv’ and ‘tf’ functions will be useful.

(c) The initial value theorem states that

$$h(0) = \lim_{s \to \infty} sH(s)$$

Calculate h(0). Does your plot confirm this calculation?

(d) The final value theorem states that

$$\lim_{t \to \infty} h(t) = \lim_{s \to 0} sH(s)$$

Calculate the final steady-state value of h(t). Does your plot confirm this calculation?

(e) Use the ‘ilaplace’ function to find the impulse response h(t) of the system. Why is the answer you get correct (look at the different terms in h(t) and relate them to the poles of H(s))?

Question 2 Convolution

An LTI system has an impulse response h(t) = u(t) – u(t-1) where u(t) is the unit-step. The input to the system is x(t) = e^{-2t} u(t).

(a) Is the system causal? Justify your answer.
(b) Is the system BIBO stable? Justify your answer.
(c) Find the output y(t) by doing a graphical convolution by hand.
(d) Find H(s) and its region of convergence by hand.
(e) Find X(s) and its region of convergence by hand.
(f) Find Y(s) = X(s)H(s) and its region of convergence by hand. Invert Y(s) to get y(t) by hand. Do you get the same answer as in part (c).
(g) Repeat part (f) using MATLAB. You should get the same answer as part (f). The 'laplace', 'ilaplace' and 'heaviside' functions will be useful. Plot the output. The function 'ezplot' will be helpful.

Question 3 Laplace Transform and Inverse Laplace Transform

(Refer Notes in Chapter 3)

1. Use MATLAB symbolic computation to find the Laplace transform of a real exponential, x(t) = e^{-t} u(t), and of x(t) modulated by a cosine or y(t) = e^{-t} cos(10t) u(t). Plot the signals and the poles and zeros of their Laplace transforms.
2. Find causal signal x(t) using MATLAB with

$$X(s) = \frac{2s+3}{s^2+2s+4}$$

Question 4 Linear Prediction of Speech

Speech processing comes under the field of discrete or digital signal processing. Although the mathematics of discrete signals has not been studied yet, the exercises below will give you a conceptual feel for some basic properties of a speech signal. The following files are emailed to you:

1. dev_a.dat – file of the sustained vowel 'a' spoken by a male speaker
2. lab3q4.m – file that accomplishes linear prediction of a 30 ms segment of the sustained vowel.

(a) Load the file dev_a.dat and plot it. Listen to it using the soundsc command. The file has 8000 samples which correspond to 1 second of speech.
(b) How many samples correspond to 30 ms of speech? Plot any 30 ms segment (or frame) of the speech signal (need not start at the first sample). You will observe a quasiperiodicity (not perfectly periodic) of the signal due to the pitch structure which is in turn due to the vibration of the vocal cords. Also, note that the pitch is the time period between successive peaks or local maxima in the signal. The pitch is both speaker and phoneme

(smallest contrastive unit in the sound system of a language which when combined lead to the formation of words) dependent. English has many phonemes and all vowels are phonemes. For example, the word ‘ship’ has 3 phonemes, namely, ‘sh, ‘i’ and ‘p’. Estimate the pitch period in the 30 ms frame that you plot. Find the time period between a few peaks (about 3 such estimates should be available) and average them. What is the standard deviation of the individual estimates?

(c) Autoregressive modeling or linear prediction is used for many speech processing applications including speech coding, speaker recognition and speech recognition. Every speech sample s(n) is modeled as a weighted linear combination of p previous samples. The model estimate of a speech sample s(n) is denoted by se(n) and expressed as

$$se(n) = \sum_{i=1}^{p} a(i)s(n-i)$$

The linear prediction error is given by e(n) = s(n) – se(n). Minimizing the squared error over a frame leads to the linear prediction coefficients a(i) and the linear prediction filter. The file lab3q4.m is emailed to you. The code is run as follows:

```
lab3q4(load(‘dev_a.dat’),start)
```

The variable start is the first sample of the 30 ms frame being analyzed. You can specify any value of the variable start. Upon running the code, three figures are seen:

- Figure 1 shows the poles of the linear prediction filter. For discrete systems, the criterion for stability is that the poles have a magnitude of less than 1. The location of the poles in relation to the unit circle is given.
- Figure 2 shows the frequency response of the filter which can be shown to be an approximation of the speech spectrum. The local maxima correspond to the resonant frequencies of the vocal tract which are known as the formants. Estimate the formants (in Hz) from Figure 2.
- Figure 3 shows the speech frame and the linear prediction error signal (e(n) as defined above). How do the two waveforms compare?

(d) Repeat part (c) using a starting sample that is at least 2000 samples away from that used in part (c). Compare figures 1, 2 and 3 from what was obtained in part (c) and comment. For example when comparing Figure 2 obtained from parts (c) and (d), do the formant frequencies change significantly? Why or why not?