EE 6330 Assignment 6

(Date Given: Friday, March 11, 2011)
(Due Date: 11:45am, Wednesday, March 23, 2011)

Objectives

- To perform non-stationary signal analysis using segmentation and linear prediction or auto-regressive modeling analysis.
- To apply linear predictive coding for compression of speech signals..

1. Background

In this computer assignment, the technique of linear prediction also known as autoregressive modeling (AR) will be used to compress speech signals to fewer coefficients. Speech signals are nonstationary in nature. That is, the mean and autocorrelation values of speech signals are variant with respect to translation along the time axis. As a consequence, AR modeling or Fourier-based spectral descriptions (such as the Fourier transform or the power spectral density, which are by definition not localized in time) are inappropriate for a physically (or physiologically) meaningful description. This inability of parametric modeling or conventional spectral analysis to deal with nonstationary signals calls for signal processing tools which would be able to monitor time variations of statistical properties of the signals and divide them into locally-stationary components: this is what *segmentation* methods are aimed at.

2. Segmentation

Segmentation of nonstationary signals can be done in two ways: fixed segmentation and adaptive segmentation. Fixed segmentation is very commonly used in speech processing. In fixed segmentation, the signal is divided into constant-length segments. An important consideration in the fixed segmentation procedure is the selection of the length of the segment (width of the window function) and the type of the window function. The segment length should be such that it is short enough for the segments to be considered as quasi-stationary, yet long enough for model parameters to be predicted accurately.

3. Linear Prediction

The linear prediction approach and AR methods are closely interrelated. If a signal $s(n)$ is produced by an AR process, the AR parameters can be estimated on the basis of the linear prediction method.

From figure 1, the AR process $s(n)$ can be written in the z-domain as

$$S(z) = G\,\frac{U(z)}{A(z)}, \tag{1}$$

i.e.,

$$G\,U(z) = S(z)A(z), \tag{2}$$

where $U(z)$ is the z-transform of a "hypothetical input" (white noise process $u(n)$), and $A(z) = \frac{1}{H(z)}$, where $H(z)$ is the transfer function of the AR model (process generator), $A(z)$ is known as the inverse whitening filter, and is the transfer function of the AR process analyzer.

Equation 2 can be written in the time domain as the convolution

$$G\,u(n) = \sum_{k=0}^{P} a_k\, s(n-k), \tag{3}$$

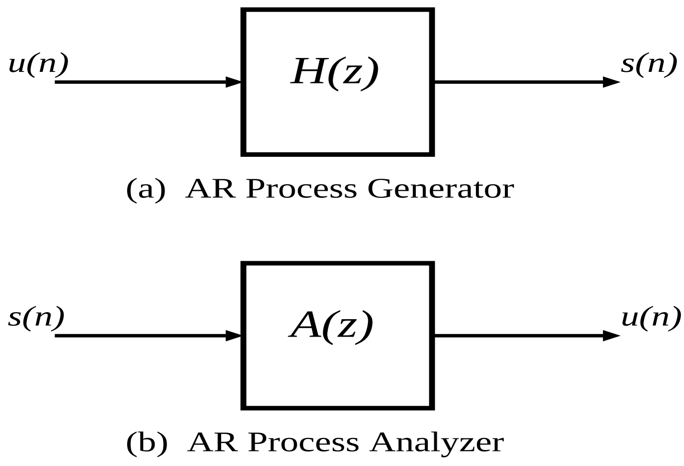

Figure 1: AR Modeling.

where a_k represents the AR coefficient at the kth stage, and P represents the filter order.

Equation 3 can we expanded as

$$Gu(n) = a_0 s(n) + a_1 s(n-1) + ... + a_P s(n-P). \tag{4}$$

As $a_0 = 1$, rearranging the above equation, we get

$$\begin{aligned} s(n) &= -\sum_{k=1}^{P} a_k s(n-k) + Gu(n) \\ &= \sum_{k=1}^{P} \alpha_k s(n-k) + Gu(n), \end{aligned} \tag{5}$$

where $\{\alpha_k\}$ are the linear prediction coefficients. Therefore, the AR model parameters of a signal are nothing but its linear prediction coefficients negated (i.e. $a_k = -\alpha_k$).

4. Spectral Estimation

The AR parameters can be used in estimating the spectrum of the given signal. We know that the transfer function of the AR process generator is:

$$H(z) = \frac{G}{1 + \sum_{k=1}^{P} a_k z^{-k}}. \tag{6}$$

The AR model spectrum can then be given as:

$$\left|H(e^{j\omega})\right|^2 = \frac{G^2}{\left|1 + \sum_{k=1}^{P} a_k e^{-j\omega k}\right|^2}. \tag{7}$$

5. Dominant Poles

By factorizing the denominator, Eq. 6 can be rewritten as

$$H(z) = \frac{G}{(z - p_1)(z - p_2)(z - p_3)...(z - p_P)}, \tag{8}$$

where $p_1, p_2, ..., p_P$ are the complex poles of the model. If the model order is an even number, the poles will occur in conjugate pairs. The distance r of a pole from the origin in the complex z- plane determines its spectral bandwidth f_B as

$$f_B = \cos^{-1}\left[\frac{(1 + r^2) - 2(1 - r)}{2r}\right]. \tag{9}$$

Poles with a large r contribute to the dominant peaks in the signal spectrum. The superior performance of poles in tracking the frequency or spectral behavior of a signal makes them an appropriate choice for parametric representation of signals with multi-peaked spectra, such as speech signals. The poles should also assist in associating the signal features with the physical characteristics of the signal source, such as resonant frequencies.

6. Programming Exercise

Divide the speech signal into fixed segments (justify the selection of window length and the type of window), and compute the AR parameters of each segment using the "lpc" function in Matlab: lpc(s_i, P) gives the $P + 1$ AR coefficients of the segment s_i.

Use the AR parameters to estimate the spectrum of each segment s_i. Modify eq. 7 accordingly to get a discrete spectrum. Compare the AR model spectrum with the FFT spectrum. Extract the poles of the AR model $H(z)$ by using the following Matlab command: [z,p,G] = tf2zp(b,a), where b and a are the coefficients of the numerator and denominator polynomials of $H(z)$, "tf2zp" converts the transfer function to zeros (z) and poles (p), and G is the gain factor. Obtain pole diagrams using the "pzmap" function in Matlab.

7. What to be Submitted?

- A description of your observations and interpretations.
- Plots of the AR spectra and FFT spectra of the segments.
- Pole diagrams of the segments.