Homework 6 (ECE6255 Spring 2010) Grade=4/100

1. Work out Problem 5.5 in Quatieri (Hint: You only need *R(0)* and *R(1)* to use the autocorrelation method in linear prediction to estimate the parameter).

2. One proposed method for estimating the pitch period based on LPC processing is to use the autocorrelation function of the LPC error signal, *e(n)*. Recall that *e(n)* can be evaluated as $e(n) = \hat{s}(n) - \sum_{i=1}^{p} \alpha_i \hat{s}(n-i)$, where the windowed signal $\hat{s}(n) = s(n)w(n)$ is nonzero for $0 \le n \le N-1$, and zero elsewhere. Define $\alpha_0 = -1$, then we have $e(n) = -\sum_{i=0}^{p} \alpha_i \hat{s}(n-i)$.

   Define the autocorrelation of $e(n)$ as $R_e(m) = \sum_{n=-\infty}^{\infty} e(n)e(n+m)$

   show that $R_e(m) = \sum_{l=-\infty}^{\infty} R_a(l) R_{\hat{s}}(m-l)$,

   where $R_a(l)$ is the autocorrelation function of the LPC coefficients, and $R_{\hat{s}}(l)$ is the autocorrelation function of $\hat{s}(n)$.

3. Consider the difference equation: $h(n) = \sum_{k=1}^{p} \alpha_k h(n-k) + G\delta(n)$

   The autocorrelation function of $h(n)$ is defined as: $\tilde{R}(m) = \sum_{n=0}^{\infty} h(n)h(n+m)$

   a) Show that $\tilde{R}(-m) = \tilde{R}(m)$;
   b) By substituting the difference equation into the expression for $\tilde{R}(-m)$, show that

   $$\tilde{R}(m) = \sum_{k=1}^{p} \alpha_k \tilde{R}(|m-k|), \ m = 1, \dots, p.$$

4. Consider two windowed speech sequences, $x(n)$ and $\hat{x}(n)$, both defined for $0 \le n \le N-1$. We perform an LPC analysis, using the autocorrelation method, and obtain two autocorrelation sequences defined as:

$$x(n) \quad 0 \le n \le N-1 \Rightarrow R(k) = \sum_{n=0}^{N-1+k} x(n)x(n+k), \ \ 0 \le k \le p$$

$$\hat{x}(n) \quad 0 \le n \le N-1 \Rightarrow \hat{R}(k) = \sum_{n=0}^{N-1+k} \hat{x}(n)\hat{x}(n+k), \ \ 0 \le k \le p$$

From these autocorrelation functions, we solve for the predictor parameters, $\alpha = \left[\alpha_0, \alpha_1, ..., \alpha_p\right]^t$, and $\hat{\alpha} = \left[\hat{\alpha}_0, \hat{\alpha}_1, ..., \hat{\alpha}_p\right]^t$, with $\alpha_0 = \hat{\alpha}_0 = -1$.

a) Show that the new prediction error, defined as $E^{(p)} = \sum_{n=0}^{N-1+p} e^2(n) = \sum_{n=0}^{N-1+p} \left[-\sum_{i=0}^{p} \alpha_i x(n-i) \right]^2$

can be written in the form $E^{(p)} = \alpha R_\alpha \alpha^t$ where R_α is a $(p+1)$ by $(p+1)$ matrix. Determine R_α;

b) Consider passing another speech sequence $\hat{x}(n)$ through the inverse filter, defined by the new LPC coefficients, α, to get another error signal

$$\tilde{e}(n) = -\sum_{i=0}^{p} \alpha_i \hat{x}(n-i)$$

show that the mean squared error $\tilde{E}^{(p)} = \sum_{n=0}^{N-1+p} [\tilde{e}(n)]^2$ can be writtern in the form

$\tilde{E}^{(p)} = \alpha^t R_{\hat{\alpha}} \alpha$ where $R_{\hat{\alpha}}$ is a $(p+1)$ by $(p+1)$ matrix. Determine $R_{\hat{\alpha}}$;

c) We can also pass the signal $x(n)$ through the old predictor filter $\hat{\alpha}$, $\tilde{\tilde{e}}(n) = -\sum_{i=0}^{p} \hat{\alpha}_i x(n-i)$

show that the predictor error is denoted by $\tilde{\tilde{E}}^{(p)} = \hat{\alpha} R_\alpha \hat{\alpha}^t$;

d) If we form the ratios $\tilde{D} = \tilde{E}^{(p)} / E^{(p)}$ and $\tilde{\tilde{D}} = \tilde{\tilde{E}}^{(p)} / E^{(p)}$, what can we say about the range of the values of $\tilde{D}$ and $\tilde{\tilde{D}}$? [Hint: note the minimum mean squared error property of LPC analysis]

5. Write a MATLAB program to convert from a speech frame to a set of linear prediction coefficients using the autocorrelation and covariance methods. Choose a steady state vowel section, and an unvoiced section, and plot LPC spectra from the two methods along with the DFT spectrum from the Hamming window weighted frame. Use N=480, p=16 for files 1.wav and 5.wav for your testing. (Don't forget that for the covariance method you need to preserve p samples before the starting sample for computing correlations and errors).

6. **Extra Credit** (+25%, 1/100 in total): For your utterances in Problem 6 compute the log magnitude spectra and log magnitude LPC spectra obtained with the autocorrelation method for a long speech section (or even for the entire utterance) containing multiple sounds. Use the multi-panel displays shown on Slide 60 in Lecture 16-18. Explain the results.