In the Name of God
Project-1
LEAST SQUARES AUDIO AND SPEECH COMPRESSION
LINEAR PREDICTIVE CODING (LPC)
Department of Electrical Engineering - Shiraz University

Instructor: Dr. Hamed Masnadi-Shirazi

1 Introduction

Speech data is sampled as $y(n);\ i = 1, 2, ..., N$ and quantized to a low bit rate during data transmission in order to have faster transmission. Directly quantizing the speech signal, resulting in the quantized signal $y_q(n)$, will significantly degrade the quality of the speech signal.

In order to reduce the degradation due to quantization we can use an autoregressive digital filter model for the original signal $y(n)$

$$y(n) = \sum_{k=1}^{l} a(k)y(n-k) + e(n) \tag{1}$$

The signal $\sum_{k=1}^{l} a(k)y(n-k)$ is a linear prediction of $y(n)$ given the l previous samples $y(n-1), ...y(n-l)$. In this project we will use $l = 10$.

We can write $e(n)$, which is called the prediction error (or residual error), as

$$e(n) = y(n) - \sum_{k=1}^{l} a(k)y(n-k). \tag{2}$$

Given the filter coefficients $a(k)$, equation (2) lets us compute the residuals $e(n)$. Equation (1) shows us how to reconstruct (recover) the signal $y(n)$. If we have a good model and we choose the $a(k)$ correctly, then we expect that the residual linear prediction errors $e(n)$ be small and have a much smaller range of values than the original signal $y(n)$. While the original signal might have a range of values between $+1$ and -1, the residuals $e(n)$ will have a much smaller range of values between $+0.1$ and -0.1. We can reduce the effects of quantization by quantizing $e(n)$ rather than $y(n)$ and transmitting the quantized residual errors $e_q(n)$, the coefficients for the model $a(k)$ and the quantization rate. The receiver can then reconstruct an estimate of the signal $\hat{y}(n)$ as

$$\hat{y}(n) = \sum_{k=1}^{l} a(k)\hat{y}(n-k) + e_q(n). \tag{3}$$

We expect that the mean squared error (MSE) found by directly quantizing the signal

$$MSE_{y_q} = \sum_{n=1}^{N} (y(n) - y_q(n))^2 \tag{4}$$

be larger than the mean squared error (MSE) found by quantizing the residuals

$$MSE_{\hat{y}} = \sum_{n=1}^{N} (y(n) - \hat{y}(n))^2. \tag{5}$$

Since the entire speech signal will contain many different sounds, the model of equation (1) cannot be used to predict the entire speech signal so we have to divide the entire speech signal into smaller sections of speech (blocks of data), that hopefully contain similar sounds, and model each block of data separately. This means that a different set of $a(k)$ model parameters have to be found for each block of data separately. Obviously, the signal will be reconstructed by the receiver in blocks as well and these blocks will be re-attached to constitute the entire reconstructed signal. In this project you are to divide the entire signal into blocks of 160 samples each.

2 General Assignment

In this project you are to do the following. For a given speech signal $y(n)$ and quantization level you need to first directly quantize the signal to find $y_q(n)$ and the corresponding mean squared error MSE_{y_q}. You must then find the least squares estimate of the parameters $a(k)$ for each block of data, then use equation (2) to find the residuals $e(n)$ and then quantize them to find the quantized residuals $e_q(n)$. You are then to use equation (3) to reconstruct an estimate of the signal for each block of data. The entire reconstructed signal and its mean squared error $MSE_{\hat{y}}$ should then be found and compared to MSE_{y_q}.

3 Matlab details

STEP 1. The sound file named failure1FIX.wav has been provided to you. Read the data from this file into Matlab and listen to it using the commands below:
[y Fs Nbits]=wavread('failure1FIX.wav');
sound(y,Fs);
Now plot the signal and include it in your report. What are the range of signal values?
STEP 2. In this step we find $y_q(n)$. For different quantization rates of $L = \{1, 2, ..., 16\}$ we can simply find the quantization intervals q by using the Matlab code:
q=(max(y)-min(y))/L;
and then find the quantized signal $y_q(n)$ using Matlab code:
yq=round(y/q)*q;
Now for each L you should listen to the quantized signal y_q and also compute the MSE_q for each L using the code below:
MSEarrayQ(i)=(y-yq)'*(y-yq)/N;
Include a plot of the MSE_q at each L in your report. How does the sound quality change as you increase L? Which L leads to acceptable sound quality? Which sounds get distorted the most? Do you think MSE is a good way to measure sound quality?
STEP 3. For each block of 160 data points, find the filter coefficients $a(k)$ $k = 1, ..., 10$ using the least squares method. So you are solving the problem of $Aa = y$ where a is the vector of $a(k)$ values, A is the matrix of previous data points and y is a vector of $y(n)$ values. The Matlab function *toeplitz* can help you easily make the A matrix. Note: Since we do not have $y(n-k)$ values for the first block, you

can assume that these values are zero.
STEP 4. You should find the residual error using the Matlab line:
error=y-A*a;
Plot the residual errors $e(n)$ and include it in your report. What is the range of $e(n)$ values?
STEP 5. Now, reconstruct the original signal $y(n)$ using equation (1). Listen to the reconstructed signal. It should be exactly the same as the original signal. Explain why this is the case?
STEP 6. Quantize the residual errors $e(n)$ using the same quantization rates and method described in STEP 2 for each block of 160 data points. Now reconstruct the signal using equation (3). For each quantization rate, listen to the reconstructed signal and also compute the $MSE_{\hat{y}}$ for each L. Include a plot of the $MSE_{\hat{y}}$ at each L in your report. How does the sound quality change as you increase L? Which L leads to acceptable sound quality? Do you think MSE is a good way to measure sound quality?
STEP 7. Compare the MSE_q plot to the $MSE_{\hat{y}}$ plot. Which method has better MSE? Which method has better sound at smaller L?
STEP 8. Repeat the steps above for a .wav file that you have either recorded yourself or one that you have downloaded from the Internet.

4 What to turn in

You should turn in a CD which includes (1) A report on your project (2) a Matlab .m file of your program (3) a .wav file from STEP 8. Your report should include your results for each STEP. You should clearly divide your report into STEPs. Failing to do so will get you a zero grade! You Matlab program should have a comment for each line and each segment of code explaining what that line or segment of code does. If you fail to comment your code, you will get a zero grade! I should be able to run your code directly from the CD and see the plots and sounds it makes. If your code has errors when I run it or does not work when I run it from the CD, you will get a zero grade!

5 CHEATING :-(

Copying a portion of someones report or Matlab code or using someones .wav file is considered cheating. You should all have different .wav files for STEP 8. If even a portion of your report or Matlab code is similar to someone else you will both get a zero on the project, be introduced to the university disciplinary committee (Komite Enzebaty) and possibly fail the course. SO DO NOT EVEN SHOW YOUR CODE OR REPORT TO ANYONE ELSE!