Alexander Smith

ECE 3522: STOCHASTIC PROCESSES IN SIGNALS AND SYSTEMS

**COMPUTER ASSIGNMENT (CA) NO. . 5: COVARIANCE AND CORRELATION**

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122

# Problem Statement

We are asked to perform some covariance and correlation calculations on our given .raw audio signal, which is sampled at 8000 Hz. Below are our instructor’s specifications.

“The tasks to be accomplished are:

1. Define a vector, **x**, of length 240 (30 msec) that contains the 240 samples of the signal starting at t = 0.9 secs. Define a second vector, **y**, which also represents 240 samples, but consists of samples shifted by k samples (e.g., k = 1 implies **y** starts one sample later than **x**). Plot the statistical correlation between **x** and **y** for k = 0, 1, ..., 512. Can you explain what you observe? (Hint: think about what would happen if the signal were a sinewave.)
2. Repeat this for t = 3.0 secs. Compare the two functions and relate them to properties of the audio signal.
3. The function you are plotting is known as the autocorrelation function. It is a minimum phase version of the actual signal. You can learn more about that in a course on digital signal processing.
4. Again, start at t = 0.9 secs. Take the first 16 samples as a vector: **x** = [x1 x2 x3 ... x16]. Compute the covariance matrix using 240 samples. Each element in the matrix is governed by the equation:
5. 
6. where i is defined over the range [0, 15] and j is defined over the range [0, 15]. Do this for t = 1.1 secs and t = 3.0 secs. Compare the two matrixes and explain why they are different.”

I used a for loop and the *corr* function to complete the first task. I had trouble with my code for the second task and was unable to finish it.



Figure 1. Correlation between x and y, with x constantly starting at sample 7200 from the speech vector and y having a starting point initially at 7200 but increasing to 7440. Both vectors are 240 samples long throughout.

For starting sample 0.9 seconds for x, as vector y is shifted in time from x, the correlation seems to get less noisy or decrease in frequency, and the amplitude is decreasing.



Figure 2. Correlation between x and y, with x constantly starting at sample 2400 from the speech vector and y having a starting point initially at 2400 but increasing to 2640. Both vectors are 240 samples long throughout.

For starting sample 0.3 seconds for x, at first it seems that as vector y is shifted in time from x, the correlation seems to get less noisy or decrease in frequency, and the amplitude is decreasing, but then it appears that it is getting noisier and louder again. This is different than when we used starting point 0.9 seconds. The two plots do appear to be smaller-phase versions of the original signal, which has a lot of frequency packed into it.

# MATLAB Code

clear;clf;close all;clc;

Faudio = fopen('rec\_01\_speech.raw','r');

spch = fread(Faudio, inf, 'int16');

fclose(Faudio);

%asked to make two vectors from spch, both 240

%samples long

% 240\*.9 / .030 = 7200 samples is the .9 sec mark

x = spch(7200:7200+240);

c=0:512;

for i = 0:1:512

 y = spch(7200+i:7200+240+i);

 c(i+1)=corr(x,y);

end

plot(c)

% 240\*.3 / .030 = 2400 samples is the .9 sec mark

x = spch(2400:2400+240);

c=0:512;

for i = 0:1:512

 y = spch(2400+i:2400+240+i);

 c(i+1)=corr(x,y);

end

figure

plot(c)

%{

x = spch(7200:7200+16);

c=0:512;

for i = 0:1:512

 y = spch(7200+i:7200+16+i);

 c(i+1)=cov(x,y,240);

end

figure

plot(c)

x = spch(7200:7200+16);

N = 240;

for n = 0:N

 for i = 0:15

 for j = 0:15

 c[i,j]=(1/N)\*sum(x(n-i)\*x(n-j));

 end

 end

end

%}

# Conclusions

This assignment demonstrated that the autocorrelation plot of a data vector will at least be reminiscent of the data vector to the viewer, and with minimum phase.