Computer Assignment (CA) No. 4: Model Fitting

Truc Le

ECE 3512: Signals – Continuous and Discrete

Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 1912

# Problem Statement

The purpose of this assignment is to continue working on the Google stock data and Speech data files, as to understand a new concept of plots called histograms. This could be accomplished by using MATLAB as our tool to observe the requested results, and these plot will helps us determine the Gaussian line of their respective bar graphs. Our aim is to understand how plotting a histogram of a certain data can produce a curve fit to determine the length of each pulse.

# Approach and Results

First, we would have to normalize all the data before we plot the histogram of both google stocks and speech file plot. After finishing normalizing both data, we would be able to execute the hist() command on MATLAB to display the histogram of the entire data. We can observe on Figure 1that the google stocks data displays the histogram plot of the entire data set. Before we draw a line of fit for this plot, we can observe that, the histogram displays all the stock values recorded on each column on the Y-axis, leaving the dates being displayed on the X-axis. The different set of colors within the bar graph represents all four columns in the stock data High, Low, Close, and Open. We can observe on Figure 2the best linear fit we managed to obtain and display on top of the google stocks data. Using the “Weibull” distribution function, which can be found on our book, Chapter 4, Hazard functions, which gives us a fair representation of the curve fit for the entire graph. Even though it was not able to cover the higher values from the stock data, it gives us a fair visualization of the distribution comparison with other plots.

Figure 3 shows the histogram plot of the speech data file. We can observe that the higher values are located within the [-0.001, 0] and [0, 0.01]. As we continue plotting a line of distribution fit, we used the normal distribution on Figure 4, to observe and compare both data distributions. Note that these values are quite similar, as they share a common mean value point near zero, and their variances almost overshadows one another.

# MATLAB Code

clear; clc;

%Google Stocks Samples & Speech File

a = xlsread('google\_v00.xlsx');

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

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

fclose(Faudio);

a2 = a/norm(a); %Normalizing data for google stock

b2 = b/norm(b); %Normalizing data for speech file

%Plot of for loop of sample for Google Stocks

figure(1)

hist(a2); %Plot histogram Google stock

title('Google Plot');

xlabel('Google Stock data');

ylabel('Google Stock value');

figure(2)

histfit(a2(:),5,'Weibull'); %Plot Weibull distribution on top of google stock

title('Google Plot + Cruve fit');

xlabel('Google Stock data');

ylabel('Google Stock value');

%Plot of for loop of sample for Speech file

figure(3)

hist(b2) %Plot histogram Speech file

title('Speech Plot');

xlabel('Speech File Data');

ylabel('Normalized data');;

figure(4)

histfit(b2,10,'normal') %Plot Normal distribution on top of Speech file

title('Speech Plot + Line fit');

xlabel('Speech File Data');

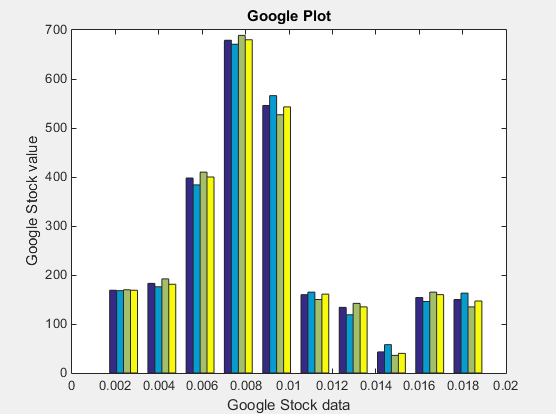
ylabel('Normalized data');

On Figure 5**,** shows code for the first part of the assignment, first we need to load and call the data values onto the workspace and normalize the data. Then we continue with using our new command hist() to plot the histogram of the google stock data. We also used a new command histfit() that takes care of plotting the distribution line of fit depending of the type of function MATLAB provides. Using the Weibull distribution, helped us observe and compare the distribution between both plots.

Figure 6shows the code content to display the histogram and linear distribution fit for the plots for the second assignment. We need to normalize all data before plotting it to the system and begin plotting the data values recorded on MATLAB to observe the histogram displayed. Using our second new command histfit() we were able to obtain the Normal distribution of the speech data file, displaying a quite similar distribution for both samples.

# Conclusions

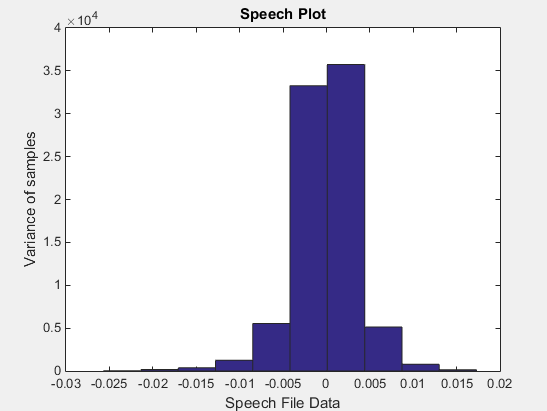
To sum up, through different distribution functions, helps us have a better understanding of the data mean and variance. Depending on each distribution used for each data, we should be able to determine the best fit for that certain plot to proceed with estimating data and analyze in depth.



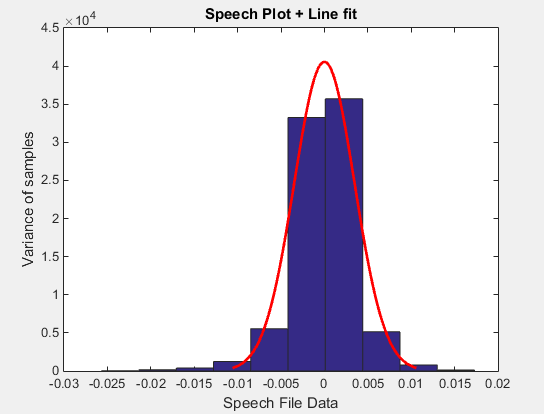
**Figure 1.** Normal histogram of google stocks data



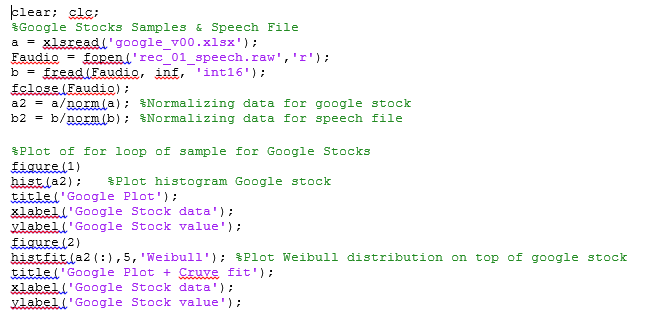
**Figure 2.** Google stock histogram **+** Weibull distribution fit



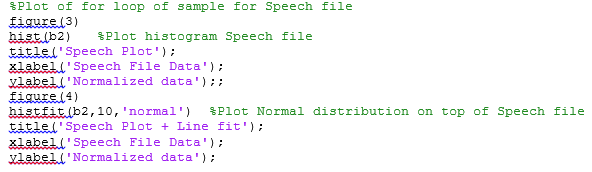
**Figure 3.** Normal histogram plot of speech data



**Figure 4.** Normal histogram of speech data + gauss fit.



**Figure 5** Code plot for google stocks



**Figure 6** Code plot for Speech file