



RESOLUCIÓN APLICADA DE PROBLEMAS: ANÁLISIS DE SEÑALES DE VOZ

Suponga que deseamos diseñar un sistema que reconozca las palabras correspondientes a los diez dígitos en inglés: "zero", "one", "two", ..., "nine". Una de las primeras cosas que podríamos hacer es analizar datos obtenidos mediante un micrófono para las diez sucesiones (o señales) correspondientes con miras a determinar si existen mediciones estadísticas que nos permitan distinguir estos dígitos.

Las funciones de análisis de datos de MATLAB nos ayudan a calcular estas mediciones con facilidad. A continuación podríamos imprimir una tabla de las mediciones y buscar aquellas que nos permitan distinguir los valores. Por ejemplo, una medición podría permitirnos reducir los posibles dígitos a tres, y otra podría servirnos para identificar el dígito específico de entre esas tres posibilidades.

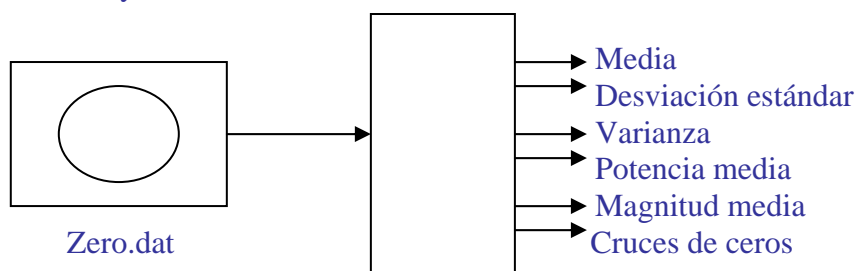
Escriba un programa MATLAB que lea y grafique un archivo de datos ASCII **zero.dat** que contenga una **pronunciación** de la palabra "zero". El programa también deberá calcular la siguiente información: media, desviación estándar, varianza, potencia media, magnitud media y número de cruces de cero. Ya hablamos de la media, la desviación estándar y la varianza. La potencia media es el promedio de los valores al cuadrado; la magnitud media es la media de los valores absolutos de los datos. El número de cruces de cero es el número de veces que se pasa de un valor negativo a uno positivo o de uno positivo a uno negativo.

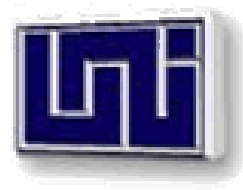
1. PLANTEAMIENTO DEL PROBLEMA

Calcule las siguientes mediciones estadísticas para una pronunciación: media, desviación estándar, varianza, potencia media, magnitud media y numero de cruces de cero además, grafique la señal.

2. DESCRIPCIÓN DE ENTRADAS / SALIDAS

El siguiente diagrama de E/S muestra el archivo que contiene la pronunciación como entrada y las diferentes medidas estadísticas como salidas.





3. EJEMPLO A MANO

Para el ejemplo a mano, suponga que una pronunciación contiene La siguiente sucesión de valores:

$$[2.5 \ 8.2 \ -1.1 \ -0.2 \ 1.5]$$

Usando una calculadora, podemos calcular los siguientes valores:

$$\begin{aligned} \text{media} &= \mu(2.5+8.2-1.1-0.2+1.5)/5 \\ &= 2.18 \end{aligned}$$

$$\begin{aligned} \text{varianza} &= ((2.5-\mu)^2+(8.2-\mu)^2+(-1.1-\mu)^2+(-0.2-\mu)^2+(1.5-\mu)^2)/4 \\ &= 13.307 \end{aligned}$$

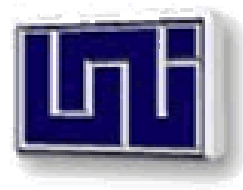
$$\begin{aligned} \text{potencia media} &= (2.5^2+8.2^2+(-1.1)^2+(-0.2)^2+1.5^2)/5 \\ &= 15.398 \end{aligned}$$

$$\begin{aligned} \text{magnitud media} &= (|2.5| + |8.2| + |-1.1| + |-0.2| + |1.5|) / 5 \\ &= 2.7 \end{aligned}$$

$$\text{numero de cruces de ceros} = 2$$

4. SOLUCIÓN MATLAB

En esta solución usamos funciones matlab para realizar la mayor parte de los cálculos. Para calcular el numero de cruces de cero, generamos un vector cuyo primer valor es $x(1)*x(2)$, cuyo segundo valor es $x(2)*x(3)$, y así sucesivamente, siendo el ultimo valor igual al producto del penúltimo dato y el ultimo dato. A continuación usamos la función `find` para determinar las posiciones de los productos que son negativos, y usamos la función `length` para contar el numero de tales productos que son negativos.



```
% este programa calcula varias estadísticas
% para una pronunciación tomada de un archivo de datos.
%
load zero.dat;
x= zero;
%
fprintf ('Estadísticas de digito \n\n')
fprintf ('media: %f \n', mean (x))
fprintf ('desviación estándar: %f \n',std(x))
fprintf ('varianza: %f \n',std(x)^2)
fprintf ('potencia media: %f \n', mean ((x.^2))
fprintf ('magnitud media: %f \n',mean (abs(x))
prod = x (1: length (x)-1).*x(2:length (x));
crossing = length (find(prod < 0));
fprintf ('cruces de cero: %.0f \n', crossings)
subplot(2,1,1), plot(x),.....
title ('pronunciación de la palabra zero'),.....
xlabel ('índice'), grid
```

5. PRUEBA

En la figura 3.13 se muestra una grafica de un archivo de datos que contiene una pronunciación de la palabra "zero". Se calculo el siguiente conjunto de valores para esta señal:

Estadística de dígitos
Media: 0.002931
Desviación estándar: 0.121763
Varianza: 0.014826
Potencia media: 0.014820
Magnitud media: 0.089753
Cruces de cero: 106

Ahora podemos ejecutar este programa con las pronunciaciones de diferentes dígitos y observar las diferencias en las medidas estadísticas. También es interesante ejecutar este programa usando pronunciaciones del mismo dígito tomadas de diferentes personas. La amplia gama de las medidas ilustradas algunos de los problemas del diseño de sistemas de reconocimiento de voz independientes del hablante. Muchos sistemas de computadora cuentan con un micrófono y altavoces. Si su computadora tiene esta capacidad., grabe la pronunciación "zero" y compare la grafica y las estadísticas con las que se muestra en esta sección.