***Aim:*** To study mathematical operation Correlation and measure degree of similarity between two signals.

***Objective:***

1. Write a function to find correlation operation.

2. Calculate correlation of a DT signals and verify the results using mathematical formulation.

3. Measure the degree of similarity using Carl’s Correlation Coefficient formula in time domain.

***Input Specifications:***

1. Length of first Signal L and signal values.

2. Length of second Signal M and signal values.

***Problem Definition:***

1. Find autocorrelation of input signal. What is the significance of value of output signal value at n=0?.

2. Find auto correlation of delayed input signal.

3. Find cross correlation of input signal and delayed input signal,

4. Find cross correlation of input signal and scaled delayed input signal.

5. Compare the resultant signals. Give your conclusion.

6. Take two input finite length DT signals and develop a function to find Carl’s Correlation Coefficient value. Determine the degree of similarity of two signals from the calculated Carl’s Correlation Coefficient value.

***Theory:***

**Correlation**

Correlation is a mathematical operation that closely resembles convolution. Correlation is basically used to compare two signals. Correlation is the measure of the degree to which two signals are similar. The correlation of two signals is divided into two ways: (i) Cross-correlation, (ii) Auto-correlation. Correlation is a bivariate analysis that measures the strengths of association between two variables and the direction of the relationship. In terms of the strength of relationship, the value of the correlation coefficient varies between +1 and -1. When the value of the correlation coefficient lies around ± 1, then it is said to be a perfect degree of association between the two variables. As the correlation coefficient value goes towards 0, the relationship between the two variables will be weaker. the direction of the relationship is simply the + (indicating a positive relationship between the variables) or - (indicating a negative relationship between the variables) sign of the correlation. Usually, in statistics, we measure four types of correlations: Pearson correlation, Kendall rank correlation, Spearman correlation, and the Point-Biserial correlation. The software below allows you to very easily conduct a correlation. Pearson r correlation: Pearson r correlation is the most widely used correlation statistic to measure the degree of the relationship between linearly related variables. For example, in the stock market, if we want to measure how two stocks are related to each other, Pearson r correlation is used to measure the degree of relationship between the two. The Point-biserial correlation is conducted with the Pearson correlation formula except that one of the variables is dichotomous. The following formula is used to calculate the Pearson r correlation:

r = Pearson r correlation coefficient

N = number of value in each data set

∑xy = sum of the products of paired scores

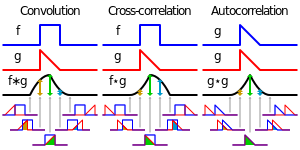
∑x = sum of x scores

∑y = sum of y scores

∑x2= sum of squared x scores

∑y2= sum of squared y scores

In signal processing, cross-correlation is a measure of similarity of two series as a function of the lag of one relative to the other. This is also known as a sliding dot product or sliding inner-product. It is commonly used for searching a long signal for a shorter, known feature. It has applications in pattern recognition, single particle analysis, electron tomography, averaging, cryptanalysis, and neurophysiology.



**Visual comparison of convolution, cross-correlation and autocorrelation.**

The cross-correlation is similar in nature to the convolution of two functions. In an autocorrelation, which is the cross-correlation of a signal with itself, there will always be a peak at a lag of zero, and its size will be the signal power. This is analogous to the distinction between autocovariance of a random vector and cross-covariance of two random vectors. One more distinction to point out is that in probability and statistics the definition of correlation Visual comparison of convolution, cross-correlation and autocorrelation.The cross-correlation is similar in nature to the convolution of two functions.In an autocorrelation, which is the cross-correlation of a signal with itself, there will always be a peak at a lag of zero, and its size will be the signal power. n always includes a standardising factor in such a way that correlations have values between −1 and +1. If and are two independent random variables with probability density functions f and g, respectively, then the probability density of the difference is formally given by the cross-correlation (in the signal-processing sense) ; however this terminology is not used in probability and statistics. In contrast, the convolution (equivalent to the cross-correlation of f(t) and g(−t) ) gives the probability density function of the sum .

**Properties**

* The cross-correlation of functions f(t) and g(t) is equivalent to the convolution of f\*(−t) and g(t). That is:
* If f is a Hermitian function, then
* If both f and g are Hermitian, then .
* Analogous to the convolution theorem, the cross-correlation satisfies where denotes the Fourier transform, and an asterisk again indicates the complex conjugate.

**Autocorrelation,** also known as serial correlation, is the correlation of a signal with itself at different points in time. Informally, it is the similarity between observations as a function of the time lag between them. It is a mathematical tool for finding repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals. Unit root processes, trend stationary processes, autoregressive processes, and moving average processes are specific forms of processes with autocorrelation.Different fields of study define autocorrelation differently, and not all of these definitions are equivalent. In some fields, the term is used interchangeably with autocovariance.

***Input:***

**Autocorrelation Program**

#include<stdio.h>

#include<graphics.h>

int m,n,x[30],h[30],y[30],i,j, k,x2[30],a[30],p,rh2[30];

void main()

{int gd=DETECT,gm,option;

int Y=0,X=200,A=0;

printf("Enter the length of the first sequence: ");

scanf("%d",&m);

printf("Enter the first sequence: ");

for(i=0;i<m;i++)

scanf("%d",&x[i]);

for(j=0;j<m;j++)

{//scanf("%d",&h[j]);

h[j]=x[j];

rh2[j]=h[-j];

printf("%d",h[j]);

}p=m+m-1;

for(i=m;i<p;i++)

{h[i]=0;}

for(i=m;i<p;i++)

{x[i]=0;

}y[0]=0;

a[0]=h[0];

for(j=1;j<p;j++) a[j]=h[p-j];

for(i=0;i<p;i++)

y[0]+=x[i]\*a[i];

for(k=1;k<p;k++)

{y[k]=0;

for(j=1;j<p;j++)

x2[j]=a[j-1];

x2[0]=a[p-1];

for(i=0;i<p;i++)

{a[i]=x2[i];

y[k]+=x[i]\*x2[i];

}}printf("The auto correlation is\n");

for(i=0;i<p;i++)

{printf("%d \t",y[i]);}

initgraph(&gd,&gm,NULL);

setbkcolor(WHITE);

line(100,300,500,300);

for(i=0;i<p;i++)

{A=y[i]\*10;

if(A>=0)

{for(j=0;j<=A;j++)

{Y=(300+j)-y[i]\*10;

putpixel(X,Y,BLACK);

}}else

{for(j=0;j>=A;j--)

{Y=(300+j)-y[i]\*10;

putpixel(X,Y,BLACK);

}}X=X+20;

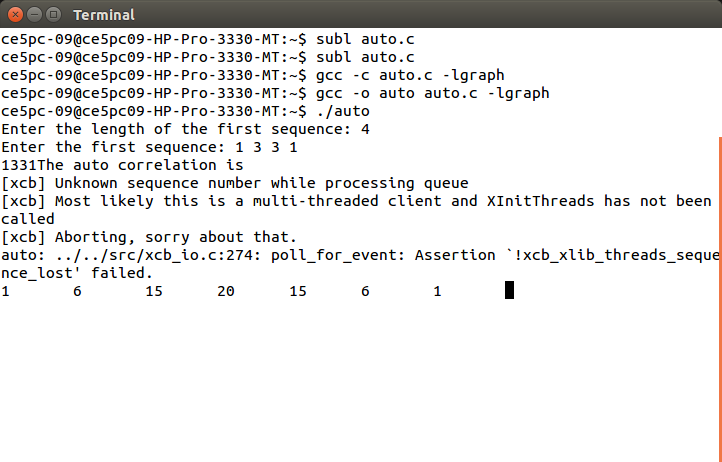
}delay(50000);

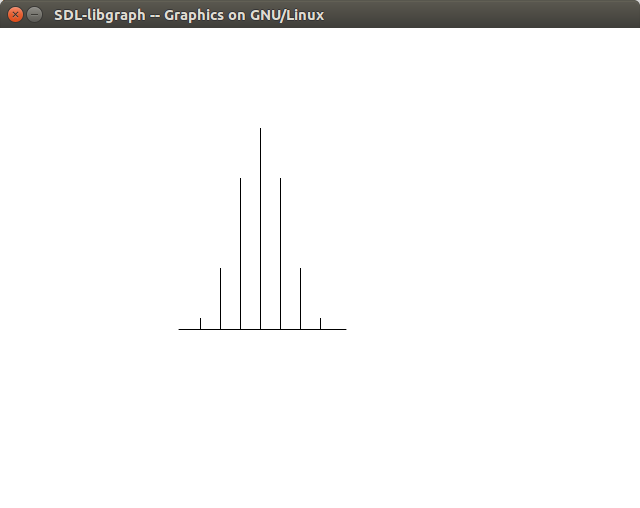
closegraph();

}

**Output:**

**Autocorrelation:**





**Cross Correlation**

#include<stdio.h>

#include<graphics.h>

int m,n,x[30],h[30],y[30],i,j, k,x2[30],a[30],p,rh2[30];

void main()

{ int gd=DETECT,gm,option;

int Y=0,X=200,A=0;

printf("Enter the length of the first sequence: ");

scanf("%d",&m);

printf("Enter the length of the second sequence: ");

scanf("%d",&n);

printf("Enter the first sequence: ");

for(i=0;i<m;i++)

scanf("%d",&x[i]);

printf("Enter the second sequence: ");

for(j=0;j<n;j++)

{scanf("%d",&h[j]);

rh2[j]=h[-j];

}p=m+n-1;

for(i=n;i<p;i++)

{h[i]=0;

}for(i=m;i<p;i++)

{x[i]=0;

}y[0]=0;

a[0]=h[0];

for(j=1;j<p;j++) /\*folding h(n) to h(-n)\*/

a[j]=h[p-j];

for(i=0;i<p;i++)

y[0]+=x[i]\*a[i];

for(k=1;k<p;k++)

{y[k]=0;

/\*circular shift\*/

for(j=1;j<p;j++)

x2[j]=a[j-1];

x2[0]=a[p-1];

for(i=0;i<p;i++)

{a[i]=x2[i];

y[k]+=x[i]\*x2[i];}}

/\*displaying the result\*/

printf("The circular convolution is\n");

for(i=0;i<p;i++)

{printf("%d \t",y[i]);}

initgraph(&gd,&gm,NULL);

setbkcolor(WHITE);

line(100,300,500,300);

for(i=0;i<p;i++)

{A=y[i]\*10;

if(A>=0)

{ for(j=0;j<=A;j++)

{ Y=(300+j)-y[i]\*10;

putpixel(X,Y,BLACK); } }

else {

for(j=0;j>=A;j--)

{ Y=(300+j)-y[i]\*10;

putpixel(X,Y,BLACK); }}

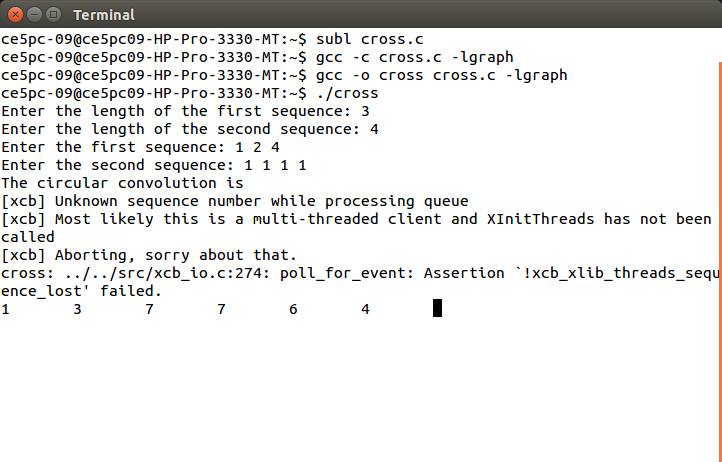
X=X+20;

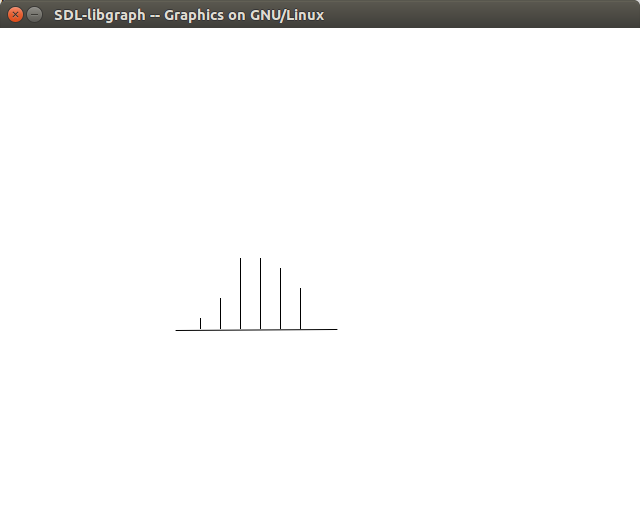
}

delay(50000);closegraph();

}

**Output:**

**Crosscorrelation:**

****

**Conclusion:**

Thus we studied & implemented mathematical operation Correlation and measure degree of similarity between two signals and achieved desired result.