Real-Time Digital Signal Processing

Prof. Rathna G N
Department of Electrical Engineering
Indian Institute of Science - Bengaluru

Lecture - 31
Correlation

Welcome back to real time digital signal processing course. So, we will discuss today correlation. To recap in the last class we have covered about overlap save method, both overlap previous to that we had covered overlap add method for continuous input signal, how we can calculate our FFT and then we had seen lab also. Coming to in today's class we will discuss about correlation. So, what is the definition of correlation or why do we need it we will look at it. So, basically a process of comparing two data sequences to obtain a measure of similarity between them.

So, most of you must be knowing that you will be trying to correlate one person with other either in terms of their habits or their look or their appearance that is what you will be looking at. In this case we will be looking with our signals whatever we are going to get the thing. So, one of it is in the application of voice recognition in audio signal processing. So, which requires comparison of different speech waveforms.

So, you are want to recognize somebody's voice and then you have stored originally and then marked it this is the original person's voice. And, some of the people their voice is recorded or from the same person, then you want to see that whether whatever stored is going to match with that particular person's voice. So, this is one of the application we need the correlation. Next, in the case of image classification in image processing which requires comparison of different image data. As I was telling, so you might have taken the image somewhere else and then you want to have it.

see that the current image whether there is any correlation or not. So, the other application one of the mostly used what we will call it in day to day life is object detection and location in basically sonar and radar systems. So, which require comparison of the transmitted signal and the signal reflected from the target objects. So, it you want to all of you know now drone is becoming mostly popular. So, if you want to identify if one of the drone is going to cross our border or whatever may be the thing.

So, you want to identify it is a foe and then you want to destroy it. So, lot of work is happening the same way in the sonar area is one of the sea area. So, submarines and other things so, they will be reflecting through the signals through radar basically and then the reflected wave which comes back you have to process it and then see that what correlation it has with the template what we have it and even you would be So, in case of aeroplanes basically, so you want to see that what kind of aeroplane which is running whether it is military or civil. So, that also reflection you will be catching through the radar and then you will be finding out. So, you know that correlation is used everywhere.

So, the fundamental measure of similarity between the two sequences we call it $x(n)$ and $y(n)$ is the sum of the products of the corresponding base of data values that is $\Sigma x(n)y(n)$ what we represented. Then, what we are going to say so, we say if the is it positively correlated? when we are going to say this if there is some kind of proportionality relationship between $x(n)$ and $y(n)$. with positive or negative values generally occurring concurrently in both sequences then some of the products will be a positive value. Because even both of them are negative we know that the product is going to give us positive and both of them are positive then it will be giving the positive value. So, then we will be saying that as a positive correlation.

When we call that as negative if there is some kind of inverse proportionality relationship between $x(n)$ and $y(n)$ with positive values in one sequence generally accompanied by negative values in the other sequence. Then we know that sum of products will be a negative value indicating that is negative correlation between the two sequences. So, when we say there is no correlation between the two sequences, if the two sequences are independent with positive values and negative values equally likely to occur in both actually the sum of products will tend to towards 0. So, we call it as due to self cancelling of the product terms in summation. So, then we say that they are not correlated.

So, continuing with the thing so, we say if two sequences $x(n)$ and $y(n)$ the cross correlation function we call it as $cxy(p)$ is defined by this equation. What do we say $cxy(p)= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} x(n)y(n-p)$. So, we say where $\pm p$ represent the number of sampling points by which $y(n)$ has been delayed or advanced in time with respect to $x(n)$. And $\frac{1}{N}$ is included as a normalization scaling factor to ensure that the cross correlation of two periodic sequences converge to the same result. as more and more sample pairs from two sequences are included in the cross correlation operation.

So, that we are trying to avoid the overflow by scaling it ok. So, how do we define the autocorrelation? When $y(n) = x(n)$, we have a special case whereby the cross correlation function becomes the autocorrelation function. The equation is given by $cxx(p)= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} x(n)y(n-p)$ in this case the maximum autocorrelation where we are going to get is when $p = 0$. Since, two identical in phase signals are being compared and the autocorrelation value decreases as p is going to increase. So, we will see how we are going to calculate our correlation as an example.

So, what we have it the same example for FFT what we have taken the thing here also we will be taking the same this thing. $x(n)$ and then $h(n)$. Here it is $x(n) = \{2,\ 0,\ 0,\ 1\}$ and $h(n) = \{4,\ 3,\ 2,\ 1\}$. So, now cross correlation by writing the table. So, n will be varying between -3, -2 -1 this is and then 0 1 2 3 to ≥ 7 we know that it is going to be 0 because we know that like linear convolution we call it as linear correlation.

Here the length is going to be L plus M minus 1. So, that is what we needed even the negative side. So, $x(n)$ we are going to write it as this thing. So, we know that in the negative region it is 0 and if nothing is mentioned we assume the starts from 0. Sometimes if you want to say that it is starting from the negative side of it.

So, you may mark with an upper arrow here stating that this is our 0th location basically. So, it is 2 0 0 1 and then elsewhere it is 0 what we have assumed p ok. And then we will be computing our correlation equation that is c_{xh} that is cross correlation with respect to p. So, first one is $h(n+3)$ because p is equal to $-$ here basically -3 it is going to be $n+3$. So, it will be 4, 3, 2, 1 what we assume the thing here and then rest of them 0s.

So, when you do the multiplication here. So, what you are going to get is. when $p = -3$. So, you are doing the multiplication $2 \times 1 = 2$ rest of them are 0s and we have to scale it by N is 4 because both the sequences are of length 4 in this case by 4 is going to give us 0.5. So, same way what you have to do is next is $n+2$. So, you will be shifting $h(n)$ sequence right by one place and then you will be computing your cross correlation here. Same way you do the thing when this thing $x(n)$ and $h(n)$ at 0 we will be seeing that is $p = 0$. So, both will be coinciding with each other that is 2 0 0 1 and 4 3 2 1 and then you will be seeing that the correlation value at this is peak it is 2.25. So, then you will be moving away from the thing p is equal to you will be giving it as 1, 2 and then 3 and then 4. So, this is our output of linear cross correlation. So, what is the difference with respect to our convolution? So, here $h(n)$ is not folded in the correlation process. So, we fold $h(n)$ in the case of convolution, here it goes as it is, it is not folded that is the difference between our convolution and then correlation. So, coming with continuing with the correlation, so what is it? it is often desirable in practice to make the autocorrelation function values independent of the signal scaling by normalizing the autocorrelation function with respect to its maximum value at zero phase.

So, that it need not have to depend on the thing. So, what is that normalization results in an autocorrelation coefficient we call it as ρ_{xx} whose values always lie in a fixed range of ± 1. So, when we take the autocorrelation $\rho_{xx}(p)$ it will be $\frac{c_{xx}(p)}{c_{xx}(0)}$. So, similarly normalizing the cross correlation function results in cross correlation coefficient ρ_{xy} which is given by $\rho_{xy}(p)$ is nothing but our $\frac{c_{xy}(p)}{\sqrt{c_{xx}(0)c_{yy}(0)}}$ at the 0 autocorrelation value what you will be taking it here for one of the signal and then autocorrelation this thing coefficient c_{yy} of the second signal.

So, and a square root you will be doing it which will normalize our this thing cross correlation function. So, the note at this places are ρ_{xy} lies in the fixed range of ±1 with + 1, 0 and -1 indicating 100 percent positive correlation. When it is 0 it is no correlation and 100 percent negative correlation respectively if it falls in this values. So, what is what does it say that our cross correlation magnitude of it should be less than or equal to square root of this function basically. As an example, we will take the same sequence $x(n)$ and then $h(n)$, we have already calculated their cross correlation coefficient.

Now, we will see the values of it that is autocorrelation $c_{xx}(0) = \frac{1}{4}\sum_{n=0}^{3} x(n)^2 = \frac{1}{4}(2^2 + 0^2 + 0^2 + 1^2) = 1.25$. And then autocorrelation of the h signal that is we have it 0 $c_{hh}(0) = \frac{1}{4}\sum_{n=0}^{3} h(n)^2 = \frac{1}{4}(4^2 + 3^2 + 2^2 + 1^2) = 7.5$. Now, we have computed cross correlation $c_{xh}(p)$ in the previous table as you can see it 0.5, 1, 1.5. So, how we can normalize this using this equation. So, what is the first value what we have is 0.5 divided by square root of are 1.25 into 7.5 will give us the value 0.16. Same way for all the autocorrelation values you can calculate are. So, normalizing the thing will give these values which have been put into table.

So, you can go and then compute for rest of them whether it is going to match or not. So, now just like our linear correlation and linear convolution. So, we have the circular correlation equivalent to circular convolution. So, how we are going to write our circular correlation in this case? We call it as $cxy(p) = \frac{1}{N}\sum_{n=0}^{N-1} x(n)y(n-p)\ for\ p = 0,1,\cdots(N-1)$. So, in this case p will also be varying between 0 to n minus 1. So, when we compute the circular this thing correlation between these two sequences as we had done it with linear correlation. So, what we have it $x(n)$ is we have 2 0 0 1 and $h(n)$ is given as 4 3 2 1. and then the N in this case is 4. So, we will be going between 0 to 3 after that it is going to repeat.

So, here when we calculate it. So, this is nothing, but 2 into 4 plus 0 into 3 plus 0 into 2 plus 1 into 1 and then scaled by 4 what we are doing it N. Then you will be seeing that it is going to give us 2.25. The next one is 1 and then next one is 1.75 and then the last one in this case is going to be 2.5 is something going to strike to you. So, you can see that what was our earlier case. So, it is 2.25 in the thing it is matching and you will be seeing only one is going to match like a linear convolution. Here also circular convolution with linear convolution there is only one value which is the at 0 which is going to match rest of them are not going to match. So, to get again our linear correlation how we are going to implement it.

So, this is one of the questions. So, how did we do linear convolution from circular convolution? By padding zeros here also you can pad zeros and then compute linear correlation from the circular correlation. So, this is going to be your assignment or take home I will call it as problem solving thing. So, you can work it out and then come back and then tell me whether you have got the results correctly or not ok. So, now, we will see some of the DFT property of our circular correlation. Just we did it for the convolution, we will do it for correlation.

So, what is it say may be stated formally with this if $cxy(p)$ is given by this equation then we call it as correlation of xy_r we will put it which is equivalent to $\frac{1}{N}(X_r Y_r^*)$. So, the multiplication is going to be X_r with Y_r^* and then scaled by $\frac{1}{N}$. So, here we say Cxy_r, X_r and Y_r are the DFTs of $cxy(p)$, $x(n)$ and $y(n)$ are the inputs and Y_r is the Y_r^* is the complex conjugate of our Y_r DFT of $y(n)$ basically. So, now we will see how we are going to compute our correlation output.

So, just like we did for the DFT. So, we will be using I will be putting it as DIFFFT. So, this is the butterfly structure what we have it. So, we have $\frac{1}{4}$ is the here also because the length of the sequences 4 what we have taken the thing. So, $\frac{1}{N}$ is $\frac{1}{4}$ in this and these are the points what we will be getting it in basically inverse of DFT basically.

So, this is a_0, a_1, b_0 and b_1. and then we will be getting out our $x(0)$, $x(2)$, $x(1)$ and $x(3)$. So, we have calculated our DFT of 2 0 0 1 in the previous class. So, which gave us $(3,\ 2+j, 1,\ 2-j)$, are the DFT of this sequence. So, I think we have not calculated DFT of $h(n)$. So, you can calculate in the same way using DITFFT that is decimation in time for fast Fourier transform butterfly structure and compute the DFT of 4, 3, 2, 1.

In this case we have assumed that we have calculated previously and using the values what we have it. So, which gives us $\{10,\ 2-j2,\ 2,\ 2+j2\}$. So, these are the DFT of 4, 3, 2, 1. Now, we will calculate the circular correlation Cxy_r what we have to calculate.

So, $\frac{1}{N}(X_r Y_r^*)$. So, if we substitute the thing. So, this is our X_r. Multiplication with Y_r^* this is the one conjugate what we have to take the thing. So, how we are going to represent this? this is $\{10,\ 2-j2,\ 2,\ 2+j2\}$. So, when you take the conjugate of this, this becomes positive and this becomes negative and then calculate these values.

So, you will be seeing that it is going to be $\{7.5,\ 0.5+j1.5,\ 0.5,\ 0.5-j1.5\}$. So, this is our value for Cxy_r. Now, how we will be calculating our $c_{xy}(0)$ as you can see the thing. So, you will be putting $a_0 + a_1$ which is nothing but $\frac{1}{4}[Cxy_0 + Cxy_2]$. So, these are the values what we have it. And then $+\frac{1}{4}[Cxy_1 + Cxy_3]$. Here we have these are the values what we have it $\frac{1}{4}[7.5+0.5]+\frac{1}{4}[(0.5+j1.5)+(0.5-j1.5)]=2.25$. So, how we calculated the circular convolution using a DFT property. So, we can calculate these values. So, the next one is $c_{xy}(1)$. So, you will be substituting these values and then calculate the thing.

So, you will be getting it as 1 in this case. Same with respect to the next two. So, you will be seeing that the DFT using DFT the circular correlation values resemble as that of the direct circular correlation what we have calculated. So, coming to the next topic we have to introduce little on random process. So, we know that real world signals such as speech, music and noise are time varying and and we know that they are random in nature. The set of all possible outcomes in any given experiment is called a sample space S.

Why do we need random process? So, you are seeing that some of the applications what we will be seeing later on. based on some of the speech and music already we have seen the thing, how a noise is introduced in the music or in the speech, how we are able to eliminate our noise if our noise is known in this case. In some of the cases noise may not be known ok, so we have to adapt to the So, which requires our random process. So, the next topic we will be covering is adaptive filter. So, we need this basic knowledge of random process only little bit of it whatever required for deriving our LMS algorithm will be covering it in this course.

So, those who want to have more they can take up as a course and then complete it. So, this is the sample space what we have defined and we have to define the random variable x. How we are going to define it? This is a function that maps all elements from sample space S into the points on the real line. So, as an example considering the outcome of our rolling fair die ok. So, we have obtained the discrete random variable that can be any one of the discrete values from 1 through 6 all of us know that our fair die has 6 values it when we roll it.

So, we do not know what we this is the outcome, but it should live be 1 through 6. So, the cumulative probability distribution function we call it as CDF of a random variable x is defined as the that basically you can see that there is a repetition here that is $F(X)=P(x\leq X)$, where capital X is real number and $P(x\leq X)$ is the probability of $(x\leq X)$.. So, we call it as probability density function of a random variable x is going to be defined with that, $f(X)=\frac{dF(X)}{dX}$.

If the derivative exists. So, two important properties of probability density function is $f(X)$ are summarized as this way that $\int_{-\infty}^{\infty} f(X)\,d(X)=1$. and $P(X_1<x\leq X_2)$ is given by our $F(X_2)-F(X_1)$. So, which is nothing but we will be putting $\int_{X_1}^{X_2} f(X)\,d(X)$. So, if

x is a discrete random variable that can be any one of this discrete That is $i = 1,2,...,$ as the result of an experiment we define the discrete probability function as $p_i = P(x = X_i)$

it belongs to this. As an example, so we consider a random variable x that has the following probability density function, $f(x) = \begin{cases} 0, & x < X_1 \ or \ x > X_2 \\ a, & X_1 \leq x \leq X_2 \end{cases}$.

So, that is we say it is uniformly distributed between X_1 and X_2 and constant value a can be computed as with this equation that is $\int_{-\infty}^{\infty} f(X)\, d(X) = \int_{X_1}^{X_2} a \cdot d(X) = a[X_2 - X_1] = 1$

as this can be seen that so the maximum value is $\frac{1}{X_2 - X_1}$. And, then we will be seeing that this is uniformly distributed between X_1 and X_2. So, $a = \frac{1}{X_2 - X_1}$ what it has been taken. So, that is the reason why we will be getting it as 1 here. So, what is it if a random variable x is equally likely to be any value between two limits X_1 and X_2 and cannot assume any value outside that range, it is uniformly distributed in the range that is $[X_1, X_2]$.

So, that is as shown in the figure it is uniformly distributed in this region then we call that is uniform density function is defined by this $f(X) = \begin{cases} \frac{1}{X_2 - X_1} & X_1 \leq x \leq X_2 \\ 0 & otherwise \end{cases}$. So, coming with some of the operations of random variables. So, the statistics what we call it as one is the first is the mean what will be defining it that is nothing, but expected value of x which is given by $\int_{-\infty}^{\infty} Xf(X)dX$. So, this is for the continuous time case is defined as this, for the discrete time case it is a $\sum_i X_i p_i$, i is varying ok. So, we represent E dot denotes the expectation operation or ensemble averaging basically.

The mean m_x defines the level about which the random process x is fluctuates. So, you will be seeing that is linear operation. The useful properties of the expectation operation are $E[\alpha] = \alpha$ and $E[\alpha x] = \alpha E[x]$ basically where α is a constant what we assume it. $E[x] = 0$, x is we call it as 0 mean random variable. So, you will be using a MATLAB function if you see the thing the *mean* calculation is given $mx = mean(x)$ computes the $mean(x)$ of all the elements in the vector x using the MATLAB function. So, as an example so, we said that our fair die rolling of a fair die.

So, we said N times that is $(N \rightarrow \infty)$. So, the what will be the probability of outcome from this. So, you will be seeing that X_i is we are taking 6 and then our probability is $1/6$ in all the cases. And in this case mean is calculated as m_x is $\sum_{i=1}^{6} p_i X_i$ So, $1/6$ of this which is going to be 3.5 and the variance the measure of spread about the mean and is defined as $\sigma_x^2 = E[(x - m_x)^2]$.

So, which is by substituting our expected value $\int_{-\infty}^{\infty}(X-m_x)^2 f(X)dX$ for the continuous time case and it is going to be sigma pi xi minus mx whole squared for the discrete time case. So, these are the mean and then variance what will be defining it, $X-m_x$ is the deviation of X from the mean value m_x, the positive square root of the variance is called the standard deviation σ_x. And MATLAB this thing function for the standard deviation calculation uses STD function ok. The various defined can be expressed as that is $\sigma_x^2 = E[(x-m_x)^2] = E[x^2 - 2xm_x + m_x^2]$. So, you are taking the expected value inside which is nothing but $E[x^2] - 2m_x E[x] + m_x^2$. So, this is the value what we will be getting it that is this becomes 0. So, it will be $E[x^2] + m_x^2$.

So, we call this is the mean square value of x. So, variance is the difference between the mean square value and the square root of the mean value. So, this is the random variable definition some of it what we will be seeing it. So, we see that if what is it mean value is equal to 0 that is $m_x = 0$,, then what happens to σ_x^2 is nothing, but $E[x^2]$. So, which we call it as P_x which is the power of x basically. Consider the uniform density function, the mean of the function can be computed by this $m_x = E[x] = \int_{-\infty}^{\infty} Xf(X)dX$.

So, we have seen the uniform which is substituted at $\frac{1}{X_2-X_1}$. So, your $\int_{X_1}^{X_2} XdX = \frac{X_2-X_1}{2}$ what will be resulted in. So, variance of the function is defined $\sigma_x^2 = E[x^2] - m_x^2$ this is what we had it $E[x^2] - m_x^2$. So, if we substitute the thing with $E[x^2]$ with that and then by simplifying, so we will be getting it as m_x^2. So, in general if x is the random variable uniformly distributed in the interval minus delta and delta. So, then we will be having $m_x = 0$ and then $\sigma_x^2 = {}^{\Delta^2}/_3$ by substituting in this you can calculate with $-\Delta$ and Δ substituting it here you will be getting ${}^{\Delta^2}/_3$ as the value.

This completes our on correlation and little bit on random process. In the next class we will be taking it up adaptive filter. So, how we are going to derive it. Thank you.