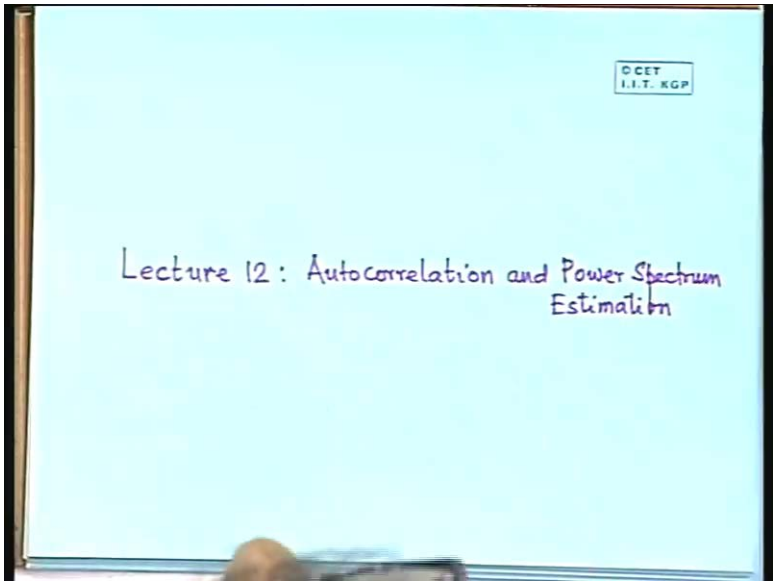
Estimation of Signals and Systems
Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture No - 12
Autocorrelation and Power Spectrum Estimation

Today's lecture, actually if you recall, we were doing estimation using linear models and then; we I thought that we have seen that, the first basic equation which is called the normal equation, involves the I mean the that equation is involves the correlation matrix. If you want to get the optimal waves you have to get the correlation matrix.

(Refer Slide Time: 0:53)

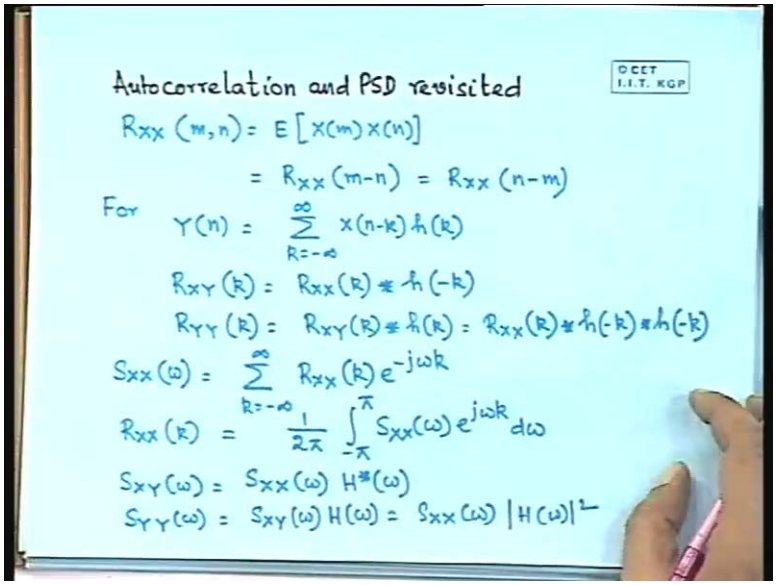


So it is a natural question as to, how am I going to get the correlation matrix? So that is why we made a you know kind of a back tracking and we today covering one topic, just to see how correlation matrix and their Fourier transforms; which are power spectral densities have to be estimated from a given data. Now basically the the problem that will be seeing is that, you know to construct process means; thus in means that there are many realizations possible, depending on the outcome of the experiment every time you get one signal. So in general auto correlation,

power spectral density, everything will mean, that you have to take an average of some quantity over ensemble, everything is an expectation. So I mean expectation is means that, it is it is an average over ensembles. Now what, why it is so, I mean that is how it is defined. But it is very difficult to to estimate that way; I mean then then we have to get multiple signals, we have to actually carry out the experiments, we have to get multiple signals then we have to carry expectation, that is the compression procedure. So in many cases, what we are interested in is that given a single realization of a process; how we can estimate those quantities and whether whether those estimates are are any good or not, because, multiple realizations are generally difficult. Secondly, theoretically speaking everything is being you know; all D F T everything is been, I mean defined from plus infinity to minus infinity and all. I mean, all realizations are are are assume to be from time, minus infinity to plus infinity. Now again that is the theoretical abstraction you cannot get that, so you will get a finite length sample sequence in general.

So what is the impact of having a finite length sample sequence; because the quantity is that we have that we want to estimate are actually defined over infinite length sequences. So the so so so that is the basic thing that we will examine today; that given a single realization of a finite length sequence, of a stochastic process, how to estimate its correlation, and power spectral density and what are the ways of improving them and things like that. That is the basic topic, so so it is you know, it is the ways of practically realizing these correlation and the power spectral density, which you can actually follow in practice. So we will just once, review the auto correlation and and P S D, this are all well known.

(Refer Slide Time: 4:05)

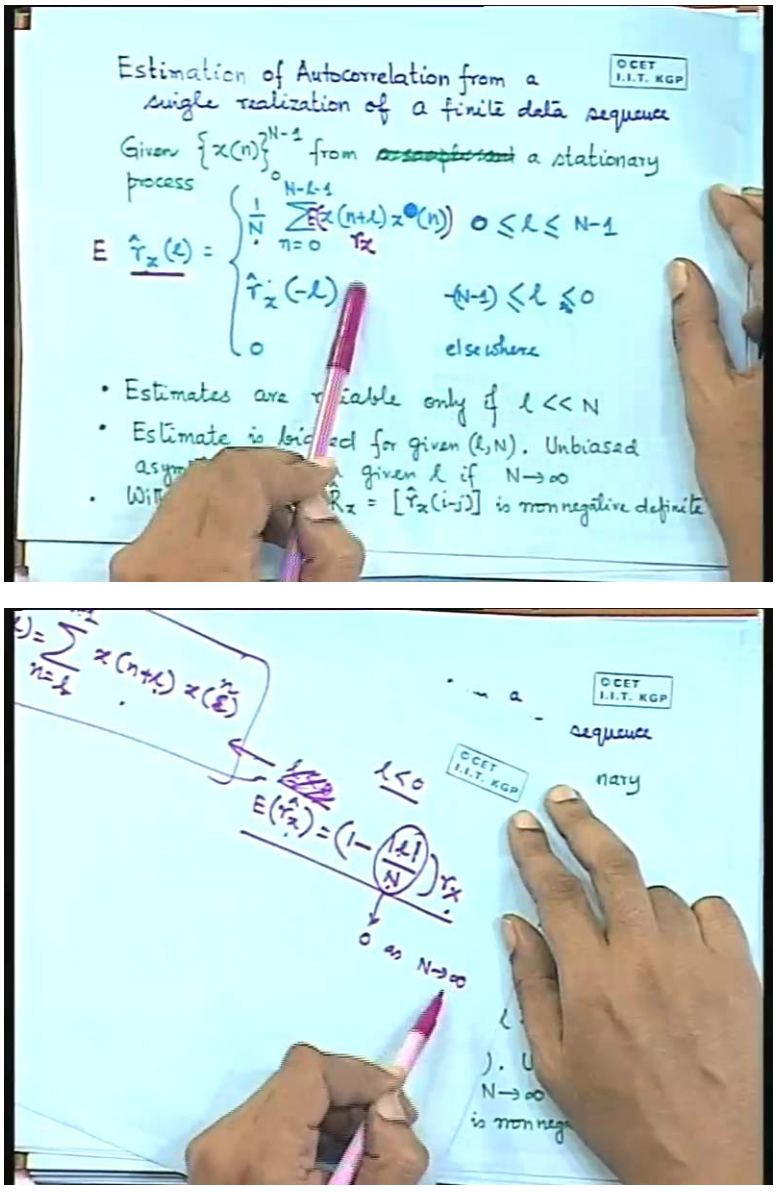


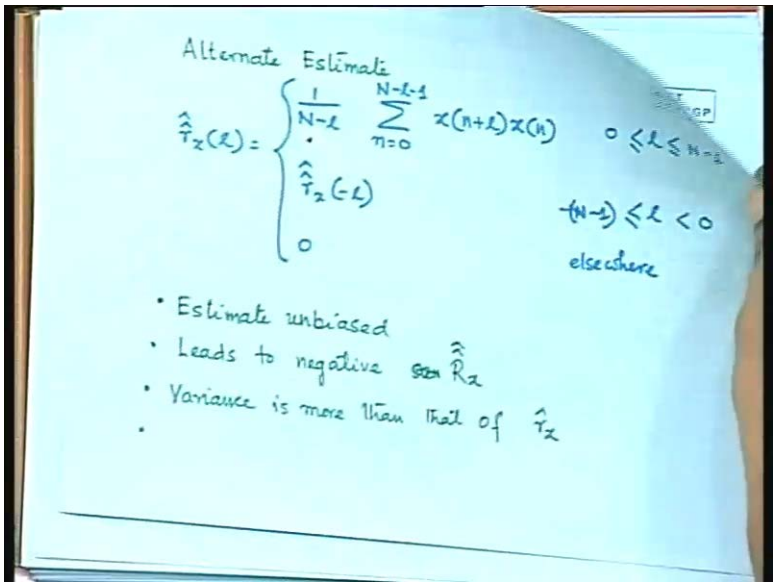
That is the auto correlation is defined as R x x m n, for a general process is expectation of x m into x n. We are talking about real processes, so there is no there is no complex conjugate involved here, everything is real. If the process happens to be stationary, that is wide sense stationary, then the correlation depends only on m minus n. And and obviously, you can exchange the arguments, I mean, you can m minus n and n minus m are the same.

So so this is the case, in case you if you can define another signal; which is obtain by a by filtering it, convolving it through a filter of impulsion response h, that is if y is the output of a filter, whose input is x, then we know that the cross correlation is defined like this. This we have already derived and the auto correlation of the output can be defined in terms of this or in terms of this. So so if this this that is you can you can want convolving them with the with the impulse response. In terms of frequency domain, the power spectral density is actually defined like this. So remember one thing that, even if see now we are talking of R x x k. So our so our time signal is is sampled, the the time sequence is discrete but that does not mean that, the frequency S x x is is discrete. It is actually a continuous function of omega, okay. That must be remembered and similarly; I if you want to get get R x x k, then you have to get it through the what is known as, the inverse Fourier transform. And the frequency domain representations of these two directions

or these two direction, that is R x x k convolving with h of minus k is is in the frequency domain like multiplication with H star omega. So finally, you get S y y omega is S x x omega into magnitude of H omega square, this relationship we have seen, okay. So now we come to an actual problem.

(Refer Slide Time: 6:26)





Our problem is to estimate, first problem is to estimate the auto correlation from a single realization of a finite data sequence. So we have a data sequence x n, which we have from time zero, from from n is equal to zero to to n is equal to capital N minus one. So we have capital N number of points and only one sequence. Now now how do you, and we have we have obviously, we assume a stationary process; so because unless you have a stationary process, you cannot write it as r x l. So it is simple that is, see we are we are assuming; what we are assuming that, the expect that the we are assuming that the that the ensemble average, can be replaced by time average. See we we have only one realization. So therefore, there is no question of having an ensemble average. We have only one realization. So if you want to make any average; we can only make a time average, because we have samples that various sequence.

So if the if the ensemble average, we are again assuming that so if the ensemble average is equal to time average, then we can take a time average. And remember that, see that these these indices have been so carefully formed, that actually actually I I I mean, ideally speaking we should form an infinite time average, but but but we cannot do that of course, because we do not have we do not have samples. So these indices have been written carefully, so that none of these will actually fall beyond zero and n minus one. See when maximum value will be n plus l, maximum value of small n is N minus l minus 1. So so so maximum value of this one will be capital N, it will not exceed, okay. And the minimum value is going to be the minimum value. So for this time

instant; this is defined as the auto correlation for positive l, and the and the autocorrelation for negative n is actually, it is simply defined it is as such defined like this. There is you you could have defined it in some other way like a like; you know so that that is for example, you could have written a different expression for this here, but then what what expression for example this is you could have written that.

Actually if you take r r x minus l, it is going to be x n minus l and x l, right. So x n minus l now, n minus l is the x n minus l into not x l, x n. This is r x l for l negative, okay. So l is less than zero, which means that; if l is less than zero, one second x l, r x l l this but for l negative, this is how r x l is defined but for l negative. So l l negative means; now this l negative means that n must belong start from l, otherwise if ends if this is if you start n from zero for l negative will go to x minus l, which is not there, that value is not there. So you could have defined it from this, to n minus 1.

This you could have defined as r x l for for l negative, you could have right. But you are not defining that, why? Because, this is not equal to this that is then r x l and and r x of minus l will not be equal in general. So so so are you following? This two summations are actually different, what I am saying is that; it it just a question of defining the summation, overall overall appropriate set of indices, such that each value is is falls within zero and n minus one. Now for positive l, if you want to impose those, then you have to put this two limits. If you want to do it for negative l, then you have to do it for this two limits. Now the moment we change the limits, the value of this sum is going to be different; because the x l, I mean there is there is nothing to guarantee, that this two sums are going to be d***, going to be same, for a for a for a given sequence. So which will immediately mean that, r x l will not be equal to r x minus l, and and this is the very fundamental quantity, I mean this is a very fundamental requirement of the auto correlation matrix.

So this we will mean that the larger auto correlation matrix will not be symmetric, all sorts of complications will arise. So therefore we will we we actually; if just deliberately define it like this that we that is we we only define the sum for for for positive l and then we say that, if l is

less than zero then, it is equal to r x of minus l. So it is always defined in terms of the positive value; so that the r x l and and and r x of minus l are are always force to be by definition, they are same, okay and it is zero elsewhere. So r x l is now defined only from minus n minus one, that is l can take value from minus n minus one to plus n minus one; simply because you have you have samples from up to minus n minus one. So therefore you cannot define correlation beyond that, if you want to define correlation beyond that; this going to be highly erroneous, actually if you if you want to define correlation beyond that, you will not get values, and and and even if l approaches close to n minus one, just imagine that the total number of terms in the sum decreases with l, right.

So finally, when when l becomes equal to n minus one, you have only one term left, what is it mean? It means that you are using a product of one term, of one realization to estimate an an expectation of a of a stochastic process, so so it is going to be highly unreliable. See you you must to be make it reliable, you must base it on more data and you must take an average over large number of data. So so therefore in general; if you I mean l should be much much smaller compare to n, only then the only then the estimate of this correlation this correlation is going to be reliable, otherwise it is variance will be too much. See the see the estimates are reliable; only if l is much much less than n, otherwise it is going to have too much variance. So that estimate will not be reliable at all.

And another important thing is that, this estimate is actually biased, for a given l and n, why? Why it is bias because if you take an expectation of this sum, what is going to be? So if you takes, see this is the single realization correlation, that you have calculated. Now let us take expectation of this, question is if I take expectation of this noun over multiple realizations, does this converge to r x? The true r x? It does not, Why? Because if you put expectation here, all this terms will be expectation, right. So each one of them will actually become r x. This is the true r x, once you have take an expectation. Now you here, you have how many terms n minus l number of terms and you are dividing by n. So you are actually having 1 minus I will put mod l, because because l can be positive and negative into r x. This is this is the expectation of r x hut, can you see this?

So you see that; for a given l and a given n expectation of r x, hut is is not r x. So even if you take several in lengths signals and to can expectation of this over over multiple signals, you will not get the true auto correlation. So the this estimate is biased for a given l and n, but obviously; this term will tend to 0 this term will tend to 0, as n tends to infinity. So it will, this estimate will become unbiased, that is it will actually converge to the true correlation provided you have taken a large number of data points, compare to l, right. This is I mean, an intuitively quite. So so this estimate is as such biased, but it is asymptotically unbiased. Asymptotically means as capital N tends to infinity; it will become unbiased but as such it is biased, okay.

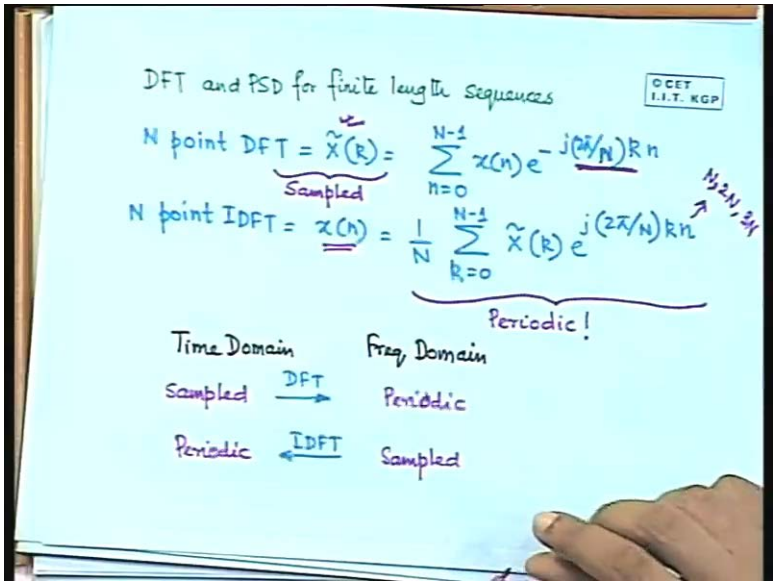
But, still these estimates are used, for example it is very easy to actually remove this bias; for example, if you are taken this N's, this estimate that is, what is the idea? The idea is that, here you have n minus l number of terms. So here why l dividing by N? Divide it by 1 by N minus l, that will make it an unbiased estimate; but there is a problem, the problem is that if you do that, then the then it it can be shown that the that the overall r x matrix will not remain positive, may not remain positive definite. r x matrix is suppose to be a positive definite matrix, that is x transpose r x must be greater than zero for all x, remember that you know earlier class, when we had talked about the normal equation, we had assume that the that the r matrix is positive definite.

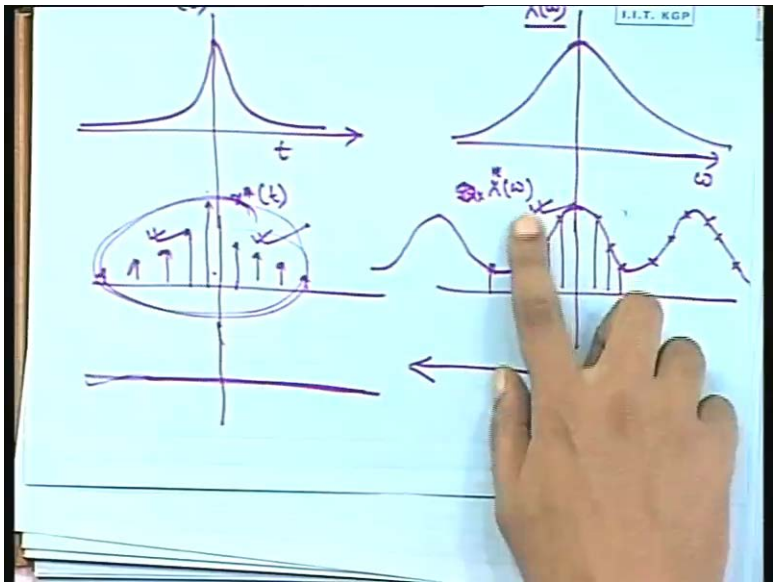
Now if you use this formula, it turns out that this matrix may not remain positive definite. Therefore; people prefer to work with this definition of correlation and keeping in mind that that, N should be sufficiently large, because otherwise if r is not positive definite, it can lead to all sorts of anonymous results, like your power spectrum. That, I mean the power spectrum that, you calculate based on this this r may be communicative, which is clearly an upset result. The power spectrum cannot go to negative. So so in the literature, mostly people will work with this definition. And so we have to only remember that; this definition will will lead to the I mean, it is a bias estimator, but the bias will go to zero as N tends to infinity. So we should take a sufficiently large value of n that is what we should remember.

Second thing is that, of course it is also I mean, this estimator variance this proofs are pretty difficult; because if you want to calculate this the the variance of these, see these itself is a is a second order moment of the samples. If you want to calculate the variance of these, then you will in generally require fourth order moments of the sample. So it it is it is quite involved, but I mean people have done it and people have found that; the between these two these are the better variance lower variance, so that is also another positive point for it just to mention, we are not going through all those calculations.

So now we come back, so now we have got a fairly intuitive formula for for calculating autocorrelation; simple time average between samples which are l apart, right. Very simple formula, actually I mean that is what anybody would do.

(Refer Slide Time: 18:52)





Now let us see that how to calculate this, the the the power spectral density from this? So we review some basic things. So here is a this a I mean, here I am discussing a very very interesting result, which those of you who were reading D S P, I mean I cannot really go into the proof and all that, but those of you who were going do doing D S P will probably, study in more much more detail. That is very interesting. That is see, suppose we have a continuous spectrum, we have a continuous spectrum of signals, something. Suppose we have some some term some time domain signal. This is x t with respect to t. Suppose, it has some spectrum, generally spectrums are, this is x omega Fourier transform of x t, this is omega. Now what will happen; if is if you sample this and and get a sample sequence? And then after sampling, if you take a take a Fourier transform, what will happen?

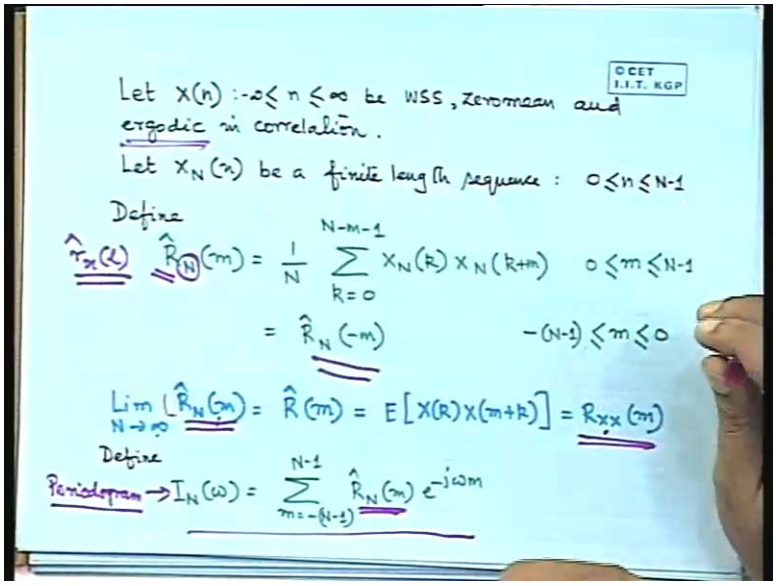
We know that x omega will now become periodic; we will get, if we sample in the time domain, this will become periodic in the frequency domain, right. This this obvious because; it is a periodic function of omega, it will become a periodic function of omega. Now the question is that, so this what this is our x x x omega, I mean this is actually x omega rather x star omega. Where x is, this is if if if this is x star t. Now the point is that, we can we can even sample this one; that will we can sample the frequency response itself, and get the value of the frequency response at certain points. Now from this sample values I can by by what is known as an inverse D F T, I can go back to to another to a to a sample time sequence, just by an inverse D F T

process; and that sample time sequence will be, a periodic version of this. So if you sample in one domain, you get a periodicity in the other domain. So if you sample in the time domain, you get a periodic thing in the frequency domain. Now if you sample this thing, you will get a periodic time function; periodic sample time function, because this was the original time function.

So if you sample this, you will get this time function periodic, right. So this is this an interesting thing because, sometimes what happens is that; I mean we must know this interpretation because in general, we will be calculating these that is when, whenever we will we will calculate the inverse DFT, we should know that what will get is actually a periodic function and between between a that is within the fundamental period, we will get the actual and other times will get same thing repeated, right. So that is what we are saying that, now here when we when we when we compute an n point D F T. Now we are computing an n point D F T that means, we have got an n point signal, based on n point signal we have we can calculate an infinite point D F T. If you take the DFT of that signal, we will actually get a continuous function, x omega. Now that we can sample at various points, so then this is this is that so so x tilda k is x n. Now you see that this is not a general omega, this is sampled at some equally distant omega point, twice pi by capital N into k. So you are sampling the frequency response, now based on this you can get back an x n sequence.

So when inverse D F T process, that is defined like this this at the. This this is the well-known N point D F T formula and this is the well-known N, N point DFT formula. Now you now you see that this thing is actually periodic in n, obviously here k is varying as an index and n is here. So so obviously, if rather than n if you have for example, for n is equal to capital N, twice N, thrice N, it has the same value. So therefore, this is periodic in N that is how you get a periodic function; this this is just an interesting observation, which we we have to remember.

(Refer Slide Time: 24:04)



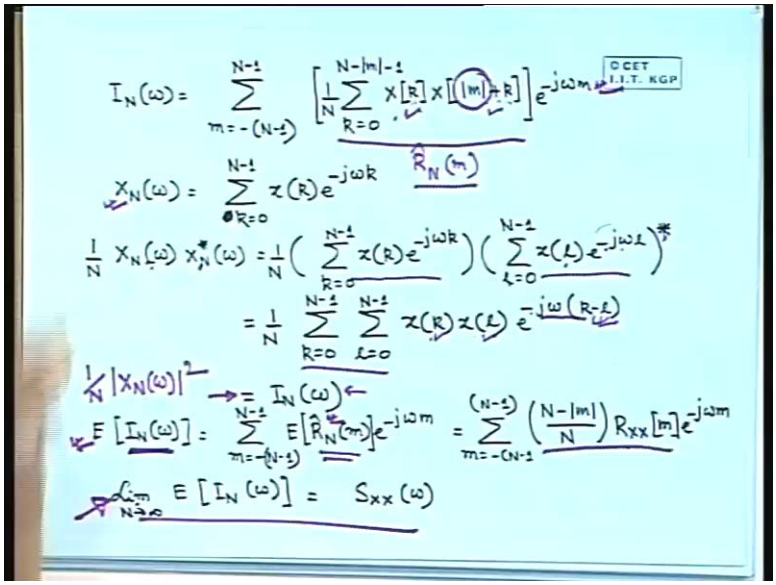
Now what is our objective? Our objective is that; we have to find out the power spectral density from a realization, there is we we we got some samples from zero to capital N and we want to calculate what is S x x omega from there. Some estimate of it, obviously we cannot calculate the actual one. So what are we going to do so first? We again define, this the the same correlation function. I have used a different notation, different parts are taken from different references. So sometimes notation is the, but but this is nothing but my same old r x l; that I had draft about, it is a same thing, okay. Only thing is that I am not put this x index here; rather than I have put an n index just to indicate that, there are n points it is a finite length sequence thing. So this is the formula, that we had see previously, same formula. And what is limit of n tend into infinity R N m, obviously; it is R x x m, limit N tend into infinity, because why, say I am not take an expectation, I am just take an N tend into infinity, because I have assume that the process is ergodic.

So is the process ergodic, then time average is equal to ensemble average by assumption. So this so this time average, long time average will will converge on the ensemble average. So therefore this will be equal to be this. If I take an expectation of this, it will it will obviously match here; whether it is ergodic or not, but I have not take an expectation. Because why I am not take an

expectation because; taking expectation is of no use to me, I want to calculate based on one realization, so when can I I? Previously, I do expectation just to show a property, but now I want to compute, so I cannot take take an ensemble average. So I one only want to stay in the time only, time average is I can use.

Now if I define a Fourier transform of these, you see what is what what is my objective? My objective is to calculate S x x omega, which is the Fourier transform of this; but I can calculate only this, so I am defining something like, something which is a which is a Fourier transform of this. Something which I can calculate and then I will see, that how could an estimate that is; incidentally this estimate is called the this estimate is known as the Periodogram. It is a very famous estimate, this is called historically, this is called the Periodogram, okay. So we will see that, how good this estimate is for a for the true S x x omega of that process. So first of all, how do we calculate this?

(Refer Slide Time: 27:05)



One One one way of calculating this is through this formula. See between the sequence I can calculate this, once I can calculate this, I can I can calculate a Fourier transform, so I can calculate this. This is one way of doing it; there is another way of doing it, so that is based on

this observation, that that 1 by N into magnitude of X N omega whole square is equal to I N omega. That is if you if you take the sequence; if you take its N point D F T, that is that is in general a complex quantity. If you take its magnitude square and then divide by 1 by N, you will get the same thing. So this is another way of, one way is to take the samples; first compute auto correlation, then takes Fourier transform, that is one way of doing it. The the other way is write in the beginning; they get the sample, compute D F T, you get X N omega, then take it's magnitude square. That will give you IN omega, that is another way, why? Because of this simple I mean, this is just an sort of an identity; that if you have I N omega, I N omega can be written like this.

This is my R hut N m, just simply substitute it, okay, and this is my X N omega. So if I multiply this X N omega magnitude square means, X N omega into X star N omega; something multiplied by its complex conjugate, will give you magnitude square. So this is X N omega, and this is again X N omega for star, only only I have written two different indices here I have use k, here I have use l. Now it is it is nothing but an a plus, b plus, c into x plus, y plus z type of thing. So you will get all ax by cz, a a ay az kind of terms.

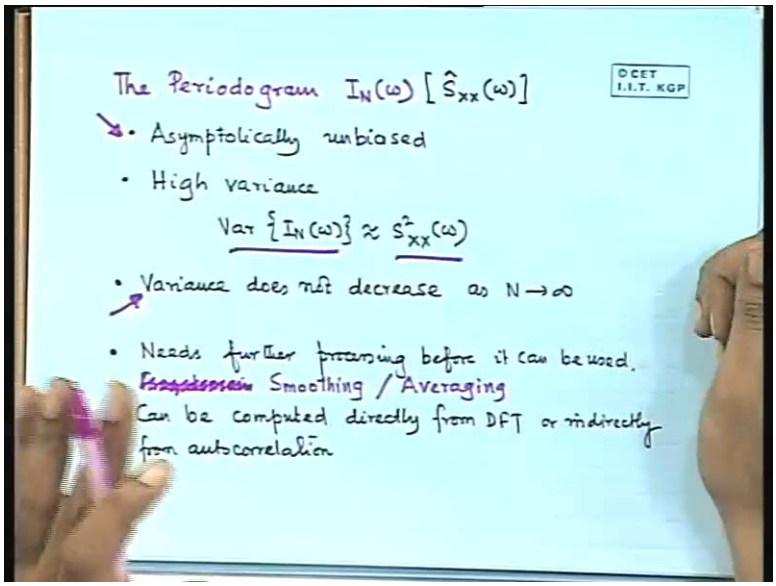
So I have now those terms, I have combined and I put it here, see this is minus g omega k, so you are getting minus g omega, k. This is minus g omega l star, see this part is real, only this part is complex. So when you take star, this will become plus; therefore minus minus plus and only this x k x l terms have been combined, very simple. So now compare this with this, see here also you get x k x l kind of terms, and here what is m? m is the difference between this index and this index. And that is appearing here; so here also same thing, two different indices and the difference between this two indices appear here. So this sums are same, so therefore; the one by N X N omega square equal to I N omega. Just nothing, just substitute and rearrange terms.

So now so so we have we have got, this two ways of calculating, the periodogram. Now so we are saying that, it is an estimate of the power spectral density, how good an estimate it is! So whenever we want to examine; how good an estimate it is, we want to see it's mean and it's variance. So it's mean is simple to estimate, it's mean is this. Simply just take expectation here,

so if you take, we already know that, the expectation of this is this, remember that, we said that it is an it is a biased estimate for a given N. So so it means that, this is also a biased estimate but it is again asymptotically unbiased; that is if N becomes too very large, then this will this will converge to R x x, and then this will converge to S x x. So again it is an asymptotically unbiased estimate, the periodogram. But but the problem is that it's variance is very large, so this is the asymptotically unbiased estimate result; that is if you compute this estimate, if you really do it for large N, then the expected value of that estimate is the true value.

So that is at least a good thing and we will see that, now expected value being large, I mean being true has gives us some reassurance. But then when we are doing for a single realization; it is not much useful, unless the variance is small and incidentally the variance is very large for this estimate. So the in so in the raw form, the periodogram estimate is a very poor estimator of S x x omega, so it is asymptotically unbiased, but if you can again compute it, in a in a very complex way and that certain assumptions of Gaussianness etcetera.

(Refer Slide Time: 32:14)



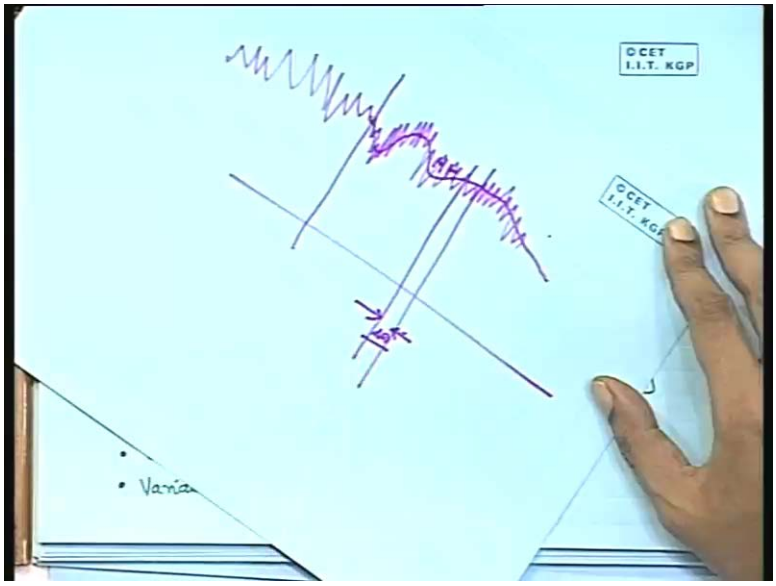
But it turns out that, the variance of this estimate is almost equal to its value. So if you are if you are want to if you want to estimate suppose some value a x and if it is variance is of the order of

a x; then you are pretty uncertain, is it not, about it? You are only certain about it, if you if you suppose, you are trying to estimate something whose value is ten; and if you are variance is say say point one, then it may be acceptable. If if you are value is ten and if the variance itself is ten, then you could be anywhere between, it could be anywhere on twenty thirty, also. So you can be I mean pretty much of from the real value, in a single case; that is why this estimate is very unreliable.

So the raw periodogram estimate is not reliable at all, though it is asymptotically unbiased. Not only that very very important thing is that; as N tends to infinity this variance does not decrease, so just by taking a long number of a large number of samples, you might you will find that the mean will converge to the to the true value. But the variance will not decrease, so you cannot do it, like that that is a that is an even bad one worse point. So therefore, it cannot be used just like this it. There has should you a, if you want to compute the periodogram estimate; you have to do further processing on it and people, see if you have some randomness which is rather large, but if you have them as kind of you know uncorrelated, then one standard way of reducing variance is by doing an average.

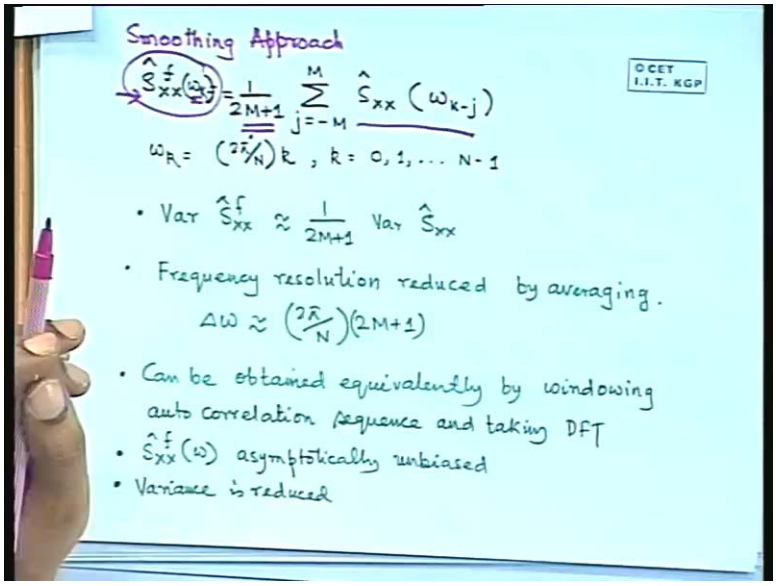
Now so there are two kinds of averaging, which are now proposed for for for periodogram estimates; to reduce their variance because, otherwise variance is too large. So those two are one is smoothing another is averaging. So what they are? That is that is that is that is very simple, so the first approach says that; suppose you got a periodogram estimate which is first approach says that, do not take suppose, you got a periodogram estimate which is like this.

(Refer Slide Time: 34:32)



On this side also, they will be symmetric. So suppose you got it like this and what it says is that; if you want to get a get an estimate smooth, it smooth it means take averages over little little frequency intervals. So if so that is like low pass filtering. So if you do low pass filtering; what will happen is that this will track it, will become like this. So the variance will decrease, but but what will happen is that, your frequency resolution; now now what you are doing is you are taking all samples of periodogram estimates over an frequency range delta omega, and you are simply taking average that is this formula.

(Refer Slide Time: 35:24)

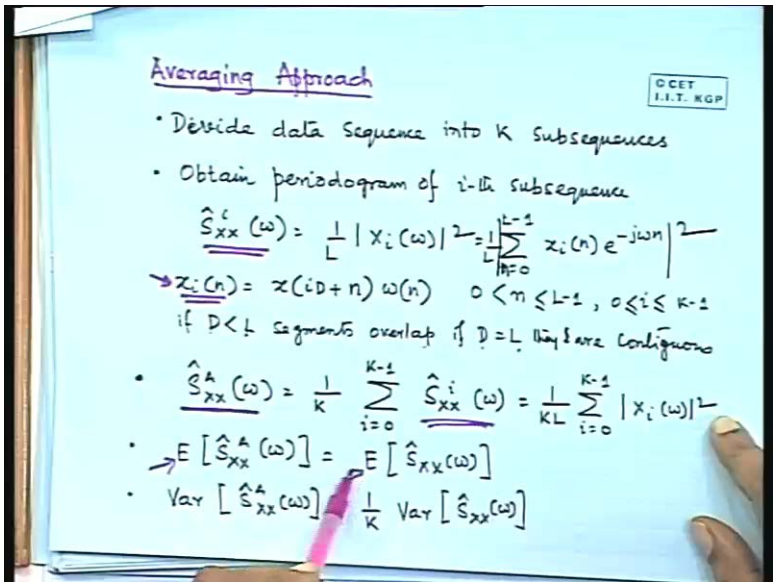


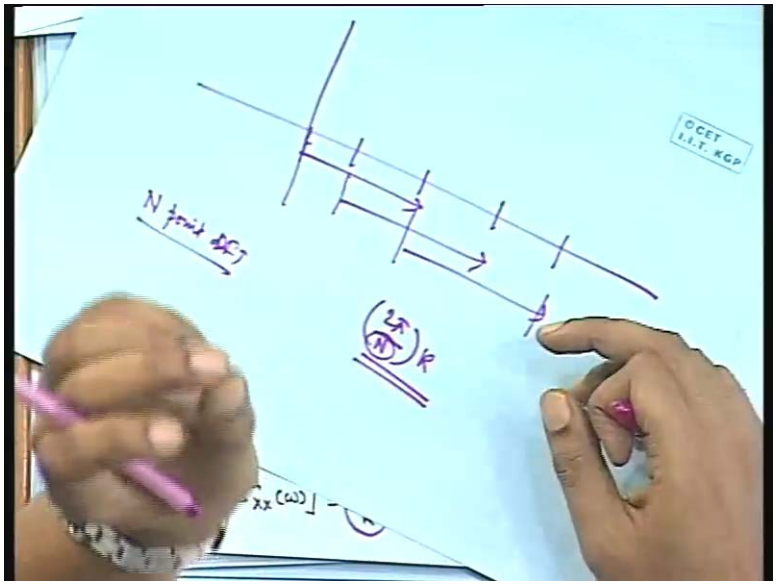
So, you are first calculating the periodogram; estimate over all samples then you are taking small small bands of frequencies, and you are calculating an average simply 1 by 2M plus 1 summed up from the same sequence. So what will happen is that, now the if the smooth sequence at every time is actually a function of the old sequence; for the last delta omega, so your frequency resolution is gone. You you do not have, you cannot catch the variations of frequencies over small small frequency, because you are going to average them. So the curve becomes smooth; so that is what you are sacrificing, so you are sacrificing on your frequency resolution to to get an estimate, which will not vary over experiments. So you are saying that okay, this may not be so accurate in frequency but; it is fairly close to what is going to happen, even if I take ten thousand such experiments. So I am increasing my reliability at the cause of at the cost of my frequency resolution.

See, something you must do, okay So so this is the what is known as the smoothing approach. And okay, finally we will see that, there how they are similar. What is the other approach; so here what you are doing is; you are taking a data sequence the whole periodogram, you are computing, after you have computed it from that periodogram, you are doing averaging on the frequency axis rather than, that you can do what is known as the sort of a sort of a virtual

ensemble averaging. That is not do an actual ensemble average, see you got a long sequence. So I will say that, no no I will chop this sequence into smaller smaller time sequences. So out of one long sequence, let me make ten small sequences; so then corresponding to those ten, I can make ten periodograms and then, I will once I calculate ten periodograms, ten different periodograms. So I will simply average them, that is that is something I can do.

(Refer Slide Time: 37:52)





So so that is the other approach, which says that, divide the divide the data sequence into K subsequences; obtain periodogram of i th subsequence, and so so you can calculate this. And then for a sequence X i n, now these can be overlapped also, it is is not that they have to be, I mean I mean I first time say that, that they are chopped, they could be chopped or they could be overlapping. That is first sequence begins, you can also do it in one overlap.. that is, if if you do not have too much data then; if you have large number of data, then you can chop it, if you do not have too much data, then you can use these two, then you can use these two then, you can use these two you can also do that, okay.

So intuitively quite simple right. So you what somehow from the data; you are trying to create, do some averaging, you are trying to create a.. situation, so that you can do some averaging; expecting that the uncorrelated things which are coming, from sample to sample by averaging, they will come down, okay. So so what you are doing is, so you can have in in general you can define X in that is the that is the given sequence like this. And if D is less than L, then this segments will overlap, and if D is equal to L, then they will be you know one after the other contiguous.

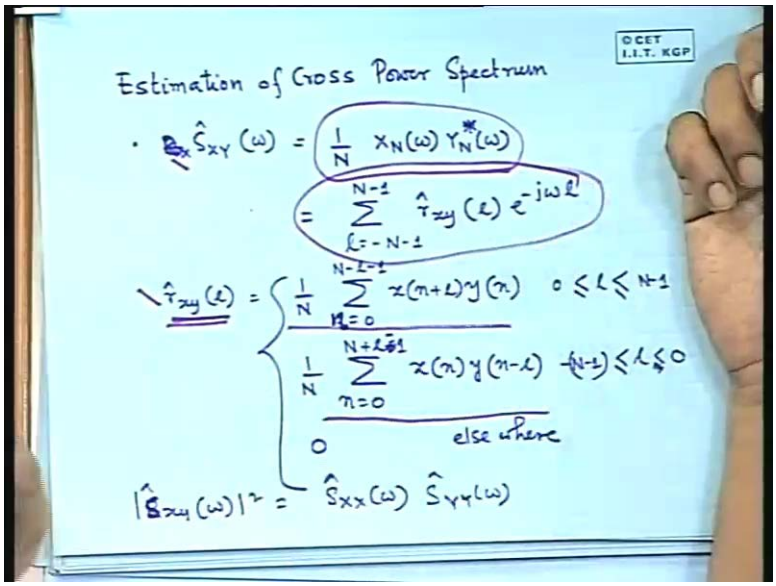
And then you can, this is the overall average its periodogram, which is simply the average of each one of the periodograms, so you sum them and divide by K, right. And it turns out that the obviously the expected value will match with the original expected value; because it is a linear operation, that you are doing and but the variance will now be one by k times. Now what is the what is the comparison between between this approaches? For example, there you were losing spectral resolution, are you not losing spectral resolution here? Definitely, you are losing. Because of the fact that, see that in the in the N point D F T, if you have an N point D F T; then then what is your what is your discrete frequency resolution? It is twice pi by N into K. These are your discrete frequency, that which you can get the D F T. So if you reduce N then you are getting larger larger. So so you are so you are not getting the D F T at finely close frequency points, if you do at ten point D F T; the D F T will the frequencies at which you will get the D F T and if you do a hundred point D F T, you will get it at much lesser frequency resolution, much finer frequency resolution.

So in the so in the second approach also; see the first approach you are losing resolution, because you are averaging, in frequency you are averaging. So it is quite obvious that, you are losing resolution. In the second approach; you are not averaging in frequency, I mean you are sample to sample, but since you have divided the data sequence into k and into k segments now, so in each segment, you now have less number of data points.

So in each segment having less number of data point means; in the in each of this periodogram frequency resolution is already gone, because you have less number of data points. So here also you are losing spectral resolution in a way, so I mean it cannot be any different, only thing is that these are two different ways of doing it, but the but finally if you have to, if with a given amount of data, if you want to have reliability, then you cannot have resolution. So there is there is certain amount of information, that you have finally, I mean actually got. So what, so you cannot by just I mean, it is not expected that; by one method you can do something miracle, so were another method. I mean if if both the methods are good enough, if you can do that, then one method is not so good.

So this is what we are doing and finally we come to we we we had known that; we we need to we are not only need this this R x x kind of quantities, we also need R x y kind of quantities and they are absolutely similar.

(Refer Slide Time: 42:12)



So we need not elaborate bad too much, you can you can so if you can calculate. So the exact way is that; you can if that you have calculated S x x, you can use exactly similar ways to calculate S x y and R hut x y. Given two sequences now X N and Y N, again over zero to N minus one, given two finite length single realizations of a stochastic process; exactly by similar formulae, for example the correlation formula will now become this, okay. And see here it is not forced like that, because R x y need not be exactly exactly symmetric. And these are so so so you can once; you can once you have this, and you can for example S x y, see the two ways of calculating; so one way of calculating S x y is take directly calculate Y N omega by D F T, directly calculate X n omega by D F T and then multiply and then have this realization. That will give you S x y omega that is the direct method. Another way is compute Rxy l and then use this formula to calculate S x y omega. That is also another way, just like the previous case.

So you can either calculate through the auto correlation or you can calculate through the through the direct D F T of the signals, both are possible. And actually you know this, if you see a D S P books, you will find that there are lots of you know, window function. There is there is something called a window function in a in in D S P; which is use to which is use to convert, finite length sequences, I mean one sequence into another. And based on the this this this properties of the window function, you get certain properties in the frequency domain. So there

are very nice window function interpretations of many of this operations. So if you are interested, you can see that but but we do not have so much time, to to carry all that.

So we will stop here today. So what we have seen is basically, what we have seen is, how to calculate auto correlation and cross correlation, how to estimate the the auto correlation and cross correlation of a stochastic process; from a finite length, single realization of that that is a I will take one get only one string is the stochastic process is infinite number of strings, depending on the outcome you get one of them. But I have only one string as measurement, that two between to only a finite number of samples from that how can I estimate the I mean the properties correlation in spectral density of the whole process. So we saw some some estimators of that and find found their properties with they are biased, unbiased variances etcetera. So this so I will stop here today.