Modern Digital Communication Techniques
Prof. Suvra Sekhar Das
G. S. Sanyal School of Telecommunication
Indian Institute of Technology, Kharagpur

Lecture - 58
Synchronization Techniques (Contd.)

Welcome to the lectures on modern digital communication techniques. So, in the previous few lectures we have been looking at the situation where channel is no longer AWGN, but it is finite band limited and we did discuss the maximum rate of signalling under such conditions, then we discussed the concept of nyquist filter, and we further stated that it is not realizable and hence we tried to investigate if there are other ways of realizing such a filter.

So, what we found is that to signal at 2 times the nyquist straight, you need to have an excess bandwidth which is the design trade off. After discussing these details what we try to do is if it was possible to restrict the signal within the nyquist bandwidth, and yet we could communicate at 2 times the nyquist bandwidth.

And we found that it is possible to do so if we could use some kind of a signalling known as partial response signalling, and we did give examples of doubinary signallings, in such a method we found that since the signals stretches more than the symbol period t. In other words because the pulse duration is longer the bandwidth is less it introduces inter symbol interference. So, you brought in the concept of controlled inter symbol interference because of controlled inter symbol interference, we could recover the signal by virtue of subtracting the interference caused; we did so by recalling that the previous time sample or the previous time instant, we could decode the symbol of interest. We moved further and we discussed 2 kinds of such pulse shapes; one is the doubinary one is the modified doubinary, and we also stated that instead of cancelling interference at the receiver one could do it at the transmitter. Meanwhile we did discuss the effect of noise and the issue of error propagation

So, we further discussed that we could do pre code; that means, at the transmitter we could do a doubinary we could do a module with 2 subtractions, by which we would get a binary sequence which would be used formodulating a binary pam. After discussing the pre coded scheme we move forward to discussed the m ary pam with such kind of signalling and we did give the expressions of how the things would work out to be and examples would be shared in the assignments and tutorials.

After looking at such signalling methods we turned our attention towards non ideal channel. In non-ideal channel we found that an impulse launched into the channel produces a response which is no longer in impulse, but it is a spreaded version of the impulse.

Now since the output of the channel has echoes and it extends beyond the impulse therefore, a sequence of symbols launched into the channel would cause inter symbol interference. And we did discussed that the overall system transfer function requires to be a raised cosine in order to make things realizable as well as inter symbol interference free. We stated that if the channel is unknown which is generally the case then the transmit filter and the receive filter are designed in such a way that when in cascade they form the raised cosine, accordingly each one of them are root raised cosine.

So, we are left with the channel as well as the equaliser. So, if the channel transfer function multiplied by the equaliser transfer function produces unity, we would get the cascaded transfer function of the transmit filter the channel the receive filter, the equaliser has raised cosine which would result in the i s i free communication

So, that results in finding or taking the equaliser coefficients as inverse of the channel transfer function. What we did see in that is if you do it in such a way wherever the channel transfer function take small values, one upon channel transfer function would take very large values. And when this factor gets multiplied with the noise we get noise enhancement. Noise enhancement decreases the performance of communication systems.

To overcome that we suggested the use of the minimum mean squared error criteria based equaliser, the equaliser look complicated, but we could break down the equaliser with certain approximations and assumptions or with certain relaxation and we found under certain special cases it turns out to be equal to 0 forcing; or in other words we all stated that the m m s equaliser restricts the enhancement of noise by including an additional parameter along with the in version of the channel coefficients. So, it does not allow the magnitude to grow very very large, it limits the amplitude growth by one upon s n r

So, after discussing the 0 forcing NMMS equaliser, we also discussed the possibility of including an adaptive equaliser; that means, instead of waiting for a long duration you can continuously adapt the tap parameters, to study that we presented 2 models. One was the system identification model in which case we would identify the channel parameters, once we identify the channel parameters you could use it to equalise the channels.

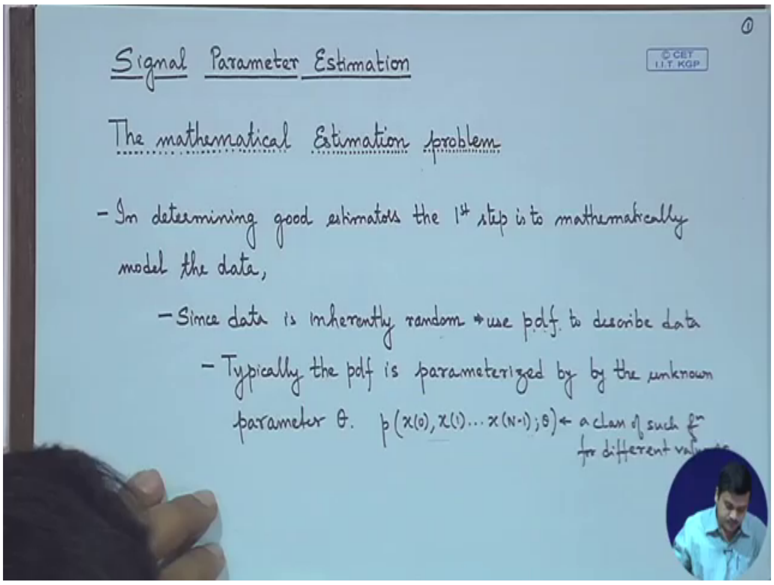
The second model we represented was of inverse system identification. In the inverse system identification one would directly estimate equaliser coefficients so that the signal out of the equaliser would match the desired signal. We also briefly stated that to estimate these parameters one should send an impulse in order to get an impulse response or find the inverse impulse response. But since sending an impulse would cause an unrealizable transmission design we stated that p n sequences generally preferred, whose transfer function would have uniform gain across the desired set of frequencies. We also stated that it is very important to design training sequences carefully because the synchronisation which we are about to start would depend on search techniques

After discussing channel equalization methods where of course, we included the method of steepest descent as well as stochastic gradient algorithm, where we briefly stated the outcome and we said that you can find out the details of it in other details subject, where as in this course we are interested only in the final expression because we are using those results in the equaliser, and again just a reminder these sections are a little bit advanced. So, from the examination point of view of this particular course will be considering the time constraint that you have in answering such questions; however, I would like to highly encourage you in pursuing those aspects if you are really interested in implementing a state of the art receiver.

So, when we move on to study the parameter estimations such as carrier frequency synchronisation and timing synchronisation, we need to resort to certain theories known as the estimation theory and detection theory. Unknowingly we have explained to you some aspects of those the theory which will repeat at certain points.

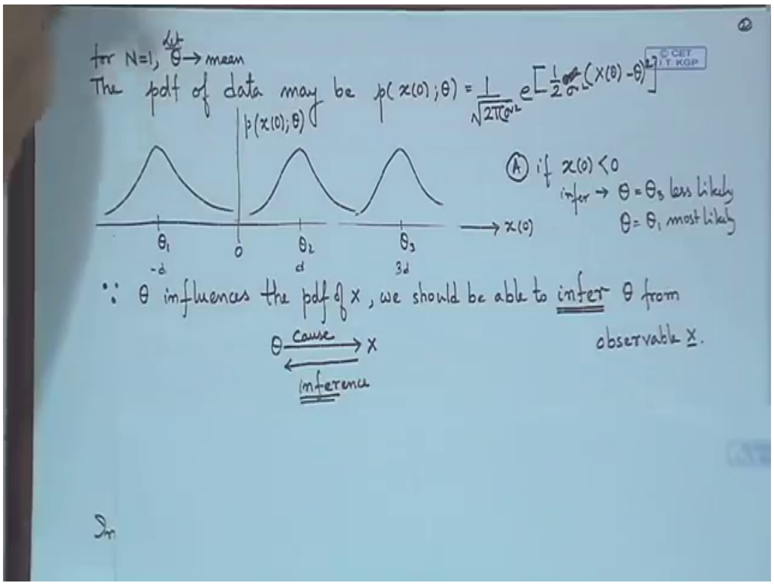
So, let us quickly brush through some of the things we discussed in the previous lecture so, that we can use them effectively in designing algorithms for carrier synchronisation and timing synchronisation.

(Refer Slide Time: 08:36)



So, we studied the parameter signal parameter estimation problem, in the signal parameter estimation problem we stated it as a mathematical statement and we said that since the data is inherently random, we use p d f to describe the data, and we also stated that in determining good estimators the first step is to mathematically model the data. Typically the p d f is parameterized by unknown parameter of let us say theta and we briefly explained with this diagram hope this will remind you of what we discussed.

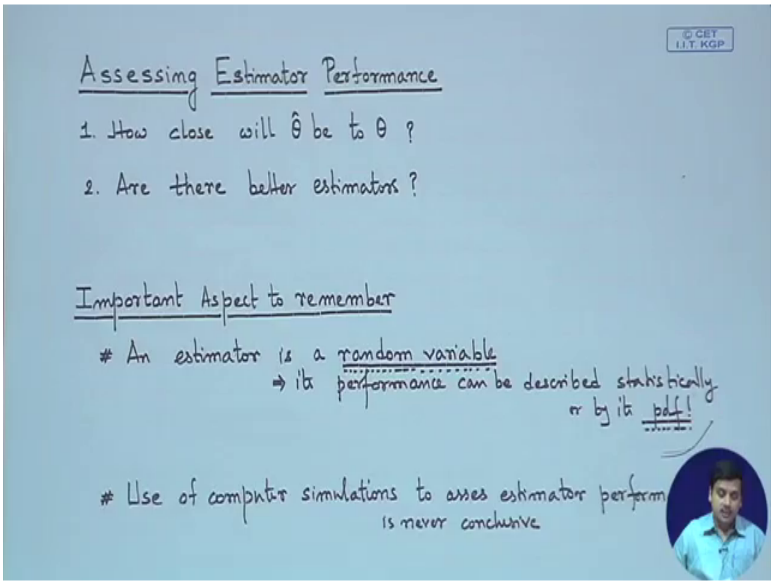
(Refer Slide Time: 09:04)



So, we did bring in the concept of likelihood function. So, this is if x 0 is observed at certain point let us say x 0 is observed here we said that if x 0 is observed here if we take go up along this curve and go to this y axis will find the likelihood of observing x using this function right which is parameterized by let us say the mean theta.

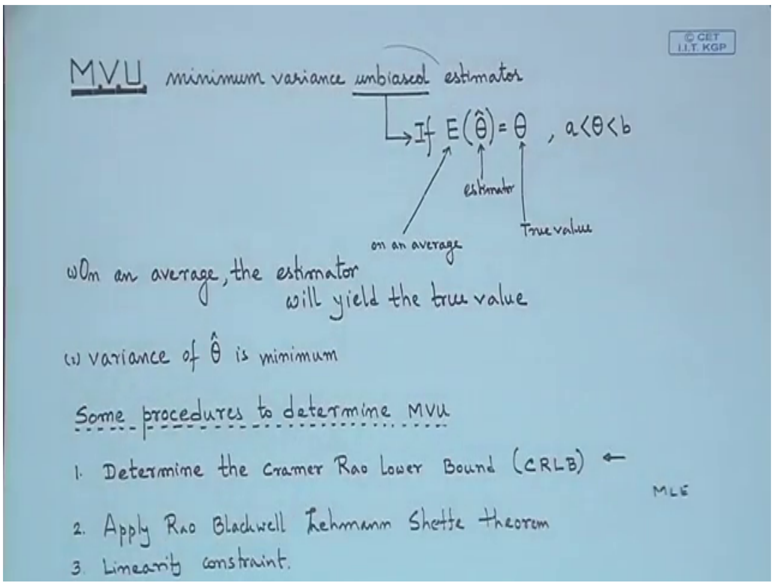
So, what we stated is that the parameter has caused the occurrence of x, and now since we have observed x, we want to infer what was the cause which has produced this x.

(Refer Slide Time: 09:59)



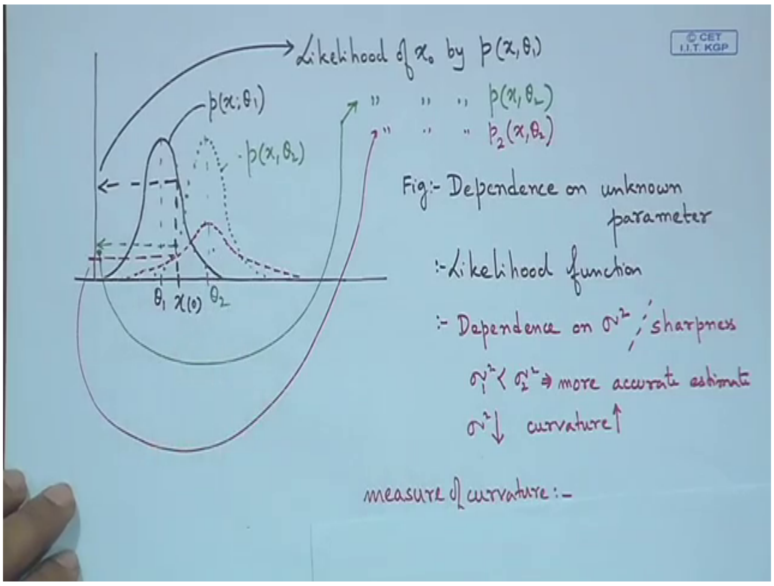
So, with this we also discussed certain other important things like when we estimate the parameter, how close will the estimate b to the actual parameter and are there better estimators.

(Refer Slide Time: 10:12)



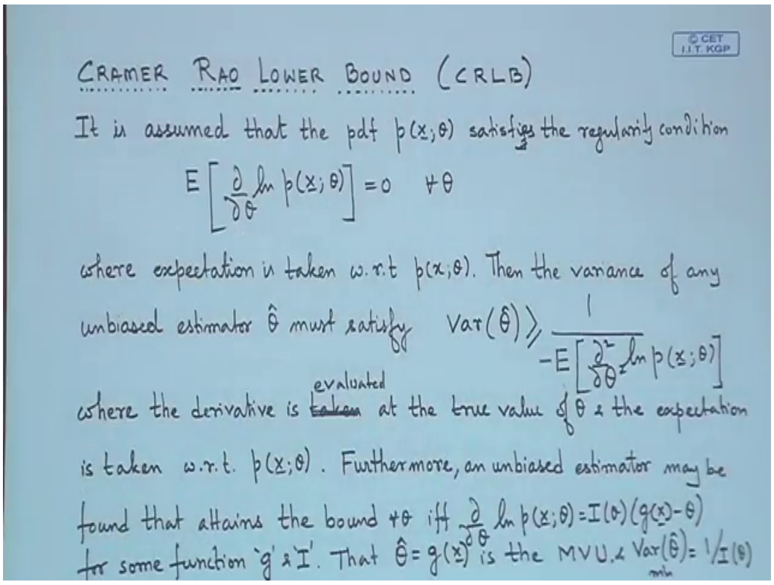
So, while discussing this we did talk about the minimum variance unbiased estimator where we said an estimatoris unbiased, if on an average it meets the value and its minimum variance if there are no other estimator whose variance is lower than the current estimator. And we discussed the cramerrao lower bound the cramerrao lower bound tells that if we I will go to that.

(Refer Slide Time: 10:34)



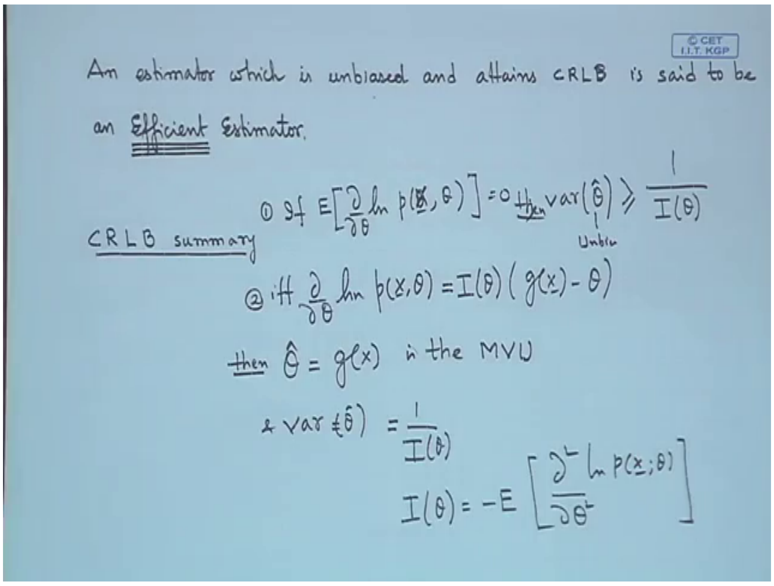
So, this picture that we looked at briefly tried to explain the likelihood function and the importance of variance with respect to the likelihood function, a smaller variance means here dependence of the unknown parameter on the p d f even more.

(Refer Slide Time: 10:53)



So, we moved further and we discussed the cramerrao lower bound, in which we stated that if the regularity condition is satisfied which has the p d f. So, we have model the data we have the p d f of the date we have assuming a p d f, we did have a short discourse on assuming the p d f with this condition is satisfied then cramerrao lower bound gives a bound on any estimated that may be found a lower bound on the variance of the estimator right. So, any estimator cannot have a variance which is lower than this that is what it stated.

(Refer Slide Time: 11:35)



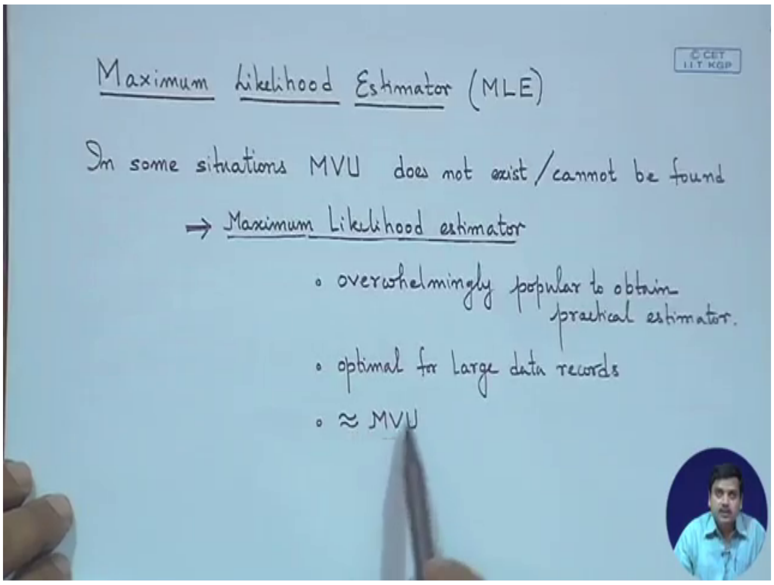
It also stated that if variable to factorise the first derivative of the log of p x, which we have written it better over here that if we are able to factorise the log likelihood function the derivative of the log likelihood function into this form then g x would form the efficient estimator; that means, it would estimate theta without any biasas well as the variance would be minimum right and the variance is given by one upon i theta where i theta is given by this term.

So, the advantage of using the cramerrao lower bound is that if you can find the first derivative of the log likelihood function, and you can factorise in the form as discussed then you will find the efficient estimator directly without much of a problem.

However the problem is we do not always get the opportunity to solve that particular equation in terms of factorising the first derivative of log likelihood function, and then we are stranded and we do not know how to find a good estimator. So, to our rescue we have the famous expression or the method which we have already used is the maximum likelihood estimator. So, since we have already used the maximum likelihood estimator the discussion in this part will almost be pretty easy, and it will be remind us about the things that we did before.

So, here we formally have the maximum likelihood estimator with us.

(Refer Slide Time : 13:11)

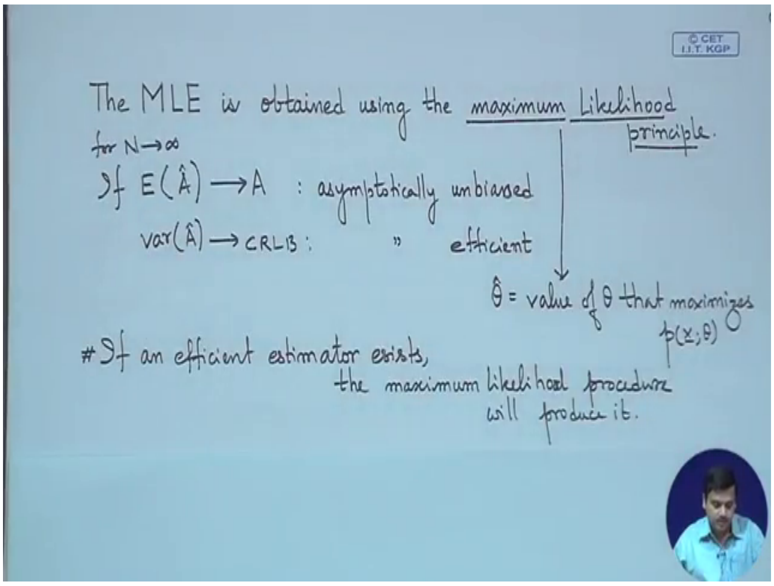


So, in some situations the MVU; that means, the minimum variance unbiased estimator either it does not exist or it cannot be found right it can happen in this situation. So, in those cases we would like to use the maximum likelihood estimator and we have already done this thing, and it is very very popular it is a very popular estimator and generally used for practical purpose. So, you must you can almost go ahead and start by using the maximum likelihood estimator, but just to remind you that when you are using the maximum likelihood estimator; that means, you have already made some assumptions and one of the strong assumptions is that, you already have the likelihood function with you.

So, if you do not have the likelihood function in that case you cannot use the maximum likelihood estimator because it depends upon the likelihood function. In those cases even this is not usable and we would resort to solutions such as the least square solution which gives us the 0 forcing algorithm or the minimum mean square algorithm as we have presented before.

So, this particular estimator is optimal for large data records and it approximately or asymptotically it is the MVU estimator we will talk more about this.

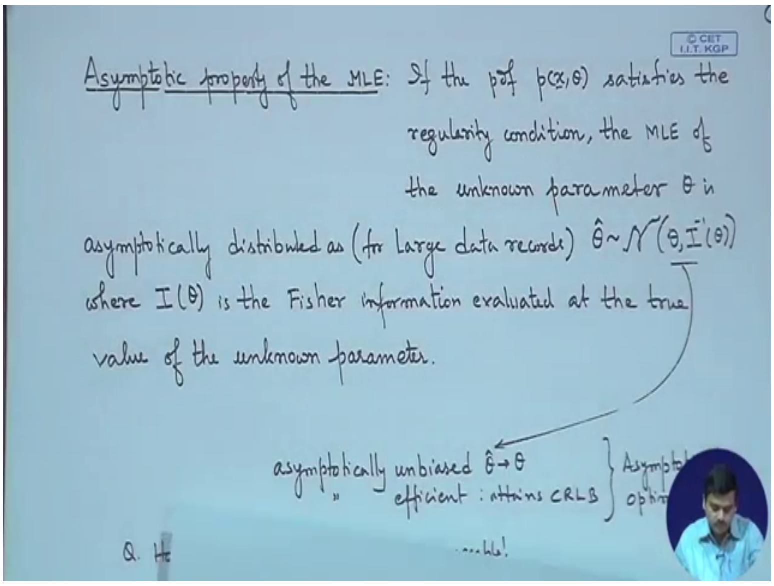
(Refer Slide Time: 14:38)



The maximum likelihood estimator is obtained using the maximum likelihood principal we have already done this, and for a very very large data set n means the number of observations, as n tends to infinity the expected value matches the data set the actual data to be estimated. And as tends to infinity the variance of the estimator tends towards the Cramer Rao lower bound so; that means, it is asymptotically unbiased and it is asymptotically efficient so; that means, if n is very large then it will yield asymptotically MVU you that is the advantage of a maximum likelihood estimator.

And the other important part is that if an efficient estimator exist the MLE will yield the efficient estimator. So, this is another very important part. So, these results tell us that if we have a likelihood function, we can almost go for the MLE method right. So, we will use the MLE method for signal parameter estimation.

(Refer Slide Time: 15:50)

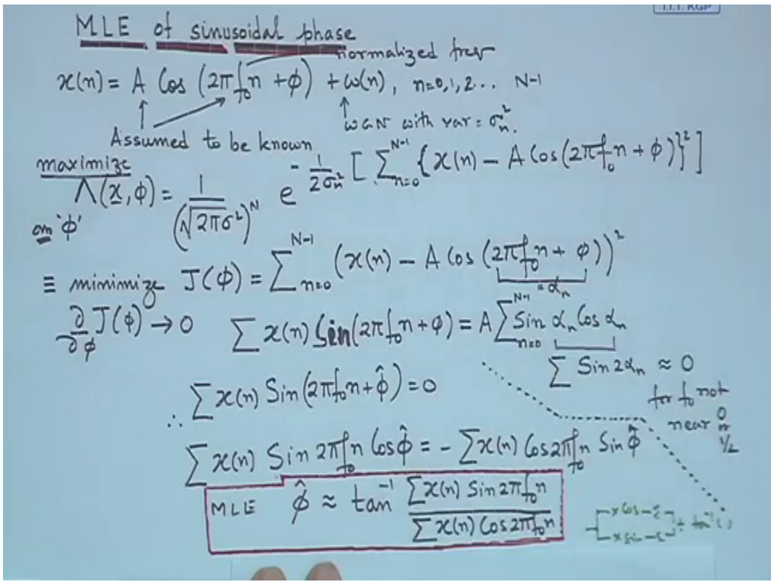


So, the another important parameter is that the estimated parameter is generally normally distributed with a mean of whatever we have said and variance as the lower bound asymptotically right ok.

One of the issues with which one may concerned with is how large in is and in many cases it is manageable. So, that is what we should sometimes, we should remember that in in quite few cases it is doable.

So, the first thing that we will discuss is since we have already studied the MLE and just to remind you that you we have already used the maximum likelihood estimator. So, in case of the signal detection, we have used the maximum likelihood principle in finding out which particular signal has been sent right. The MAP criterion uses the likelihood function, the maximum likelihood we use the maximum likelihood principal when we said equi probable transmitted signals rights. So, we have the map detector and their MLE detector. So, we use the maximum likelihood philosophy which states we use the likelihood function maximize the matrix and you have your results.

(Refer Slide Time: 17:11)



So, the first example that we will take in this case now is the maximum likelihood estimator of the sinusoidal phase. So, we have a discrete signal model x n is equal to A cos 2 pi f 0 n plus phi plus w n. So, there is noise and there is the modulated signal there is the sorry the carrier signal with some f naught, f naught is the normalised frequency and we will assume that we will know the amplitude as well as the frequency, and we are interested in finding the phase. W is white Gaussian noise with variance sigma squared n in our model it is n naught by 2.

So, now we take a look at the likelihood function. So, we would generally represent with capital lambda as a likelihood function. So, it is nothing, but the p d f where we have fed the value of x. So, if you look at this expression we have white Gaussian noise. So, this is the mean of the signal where everything is known except this parameter right. So, this will this condition on this this is the mean and then if we have x like x 1 x 2 x 3 x 4 up to x n then the joint distribution would be the product of the individual or the marginal distribution because we have white Gaussian noise, which would lead to the independence of the observes data.

So, the likelihood function of x underscore indicating a vector parameters by phi is this expression which you are pretty familiar with x n minus the mean squared and the summation, because you have a product of the p d f s. So, the summation comes we have seen this and there is a raise to the power of n because there are n such observed data.

So, instead of taking this whole function since we have to estimate phi, phi is available only here. Since phi is available only here we can focus only on this part and maximizing this means you minimise the argument, because if smaller value in the argument would be maximization of the likelihood function. So, you focus on this metric, which will remind us about the distance metric as we had done before.

So, if you take the derivative of the cost function which is in the argument and take with respect to phi, because we are maximizing with respect to phi and set it to 0. So, if you look at this you are going to get the expression because there is a cos squared term you are going to get x n squared A squared cos squared and 2a cos. So, from all those terms we want to get an expression which appears here where for brevity I have math this as alpha n 2 pi f naught n plus phi as alpha n, because of shortage of space i have used alpha n over here, and now this term would turn out to be a sin twice alpha n right and if you are taking the summation for f 0 not close to 0 or f naught; that means, not close to the extreme values of frequency this summation would turn out to be approximately 0.

So, we are making an approximation over here that for large n and f naught not close to 0 or half, we would get the right hand side of this expansion to be almost equal to 0. So, then we can say that the left hand side is equal to 0 and now we have a sin of a plus b i will term this as a plus b which will expand and they will be you will get the trigonometry expansion of this, and you are going to split the terms on both sides of the equality and you will be landing up with an expression as in here where you have a cos phi term a sin phi term which are the estimates and the sin 2 pi f naught n and cos 2 pi f nought n and x n. X n is the observed data and this is a parameter to be estimated.

So, from this one can solve as the estimate of phi as tan inverse because this is outside the summation. So, you can clearly bring it down and tan inverse of sin of this fraction summation x n sin 2 pi f naught n, upon x n cos 2 pi f f naught n. Some of the few important things to observe is that we wanted to estimate this phase using the MLE and we use the log likelihood function from the likelihood function we took the cos function, you could have set the derivative of this to 0 and you could have also got the same result, but we focused mainly on here and by making another approximation where the summation tends to 0, we have come to the expression of the estimate of phi which is the tan inverse of the observed data multiplied by sin 2 pi f naught n upon the ratio of tan inverse of the observed data times cos 2 pi f naught n.

So, what it tells is that if i observe x n I am going to multiply, I am going to get x n right. So, I have to split x n into 2 parts in one part i will multiply cos 2 pi f naught n, look at that f naught is known for us right and in this part I want to multiply sin 2 pi f naught n right n is the index, and then I can take the summation and I will feed to a device which will take if this is a this is b tan inverse of b upon a right. So, this is going to give me the phase so; that means, using the maximum likelihood estimator we can directly calculate the phase of the unknown signal and of course, one could find what kind of a behaviour does it have and so and so forth. So, why such thing is important that why do we need to find the phase of the sinusoid carrier in this kind of a situation, is something which we should see in the next lecture.

So, at least one important thing that we could conclude is that using the MLE we could directly arrive at the estimator of phi which is otherwise not clearly available from this expression right. Because if you would like to estimate phi by solving algebraic methodology you would say x n minus w n right divided by a cos inverse of that, and then take away 2 pi f naught n.

Now, we do not have this noise. So, you cannot use such a model you have to use certain methodology over here and since this can be modelled as Gaussian p d f. So, it helped us in reaching this particular estimate of phi.

So, we continue with the discussion on why this estimation of the carrier phase is important in the next lecture, and then we will proceed on to find the total carrier phase or the carrier frequency at the receiver and which is very much necessary for decoding of the signals appropriately.

Thank you.