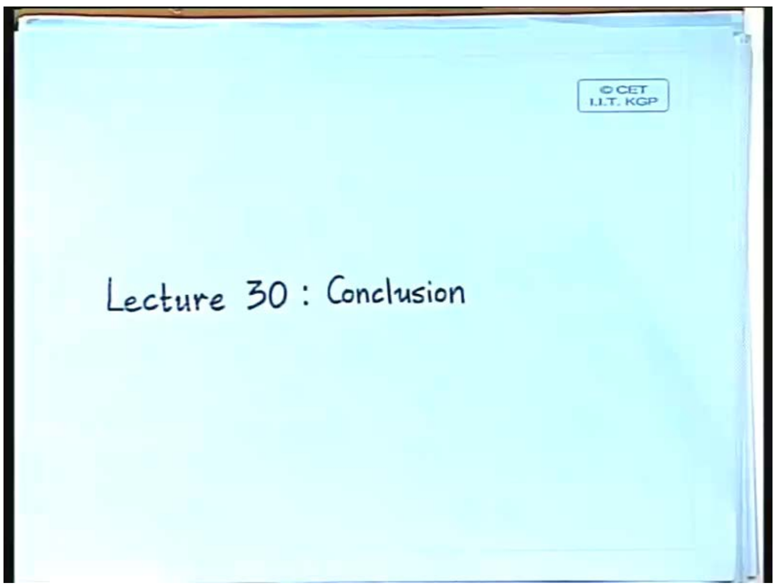
Estimation of Signals and Systems
Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 30
Conclusion

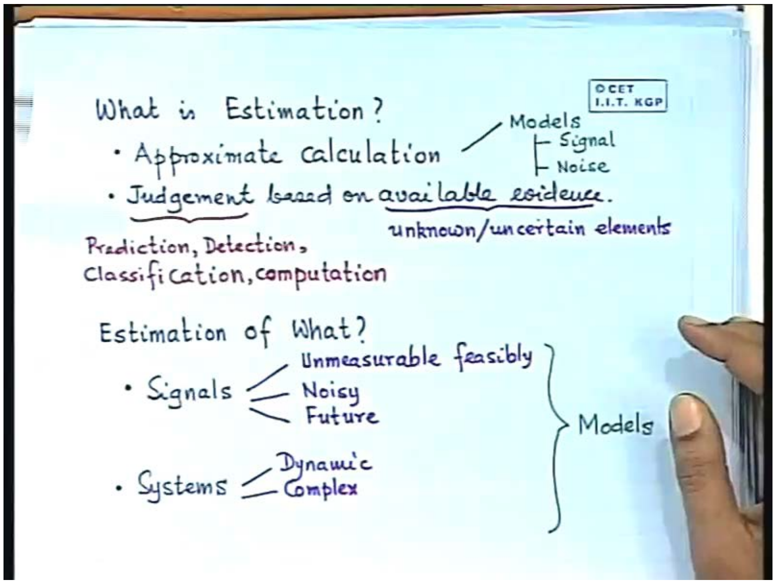
Good afternoon. So, we have come to the end of the semester's journey, which we began in July.

(Refer Slide Time: 01:13)



We started, right from probability theory, and went on to explore various major classes of estimation problems. And in the last class we have also seen that, there are many other types of estimation problems which exist. So, this class is to reflect back, on what we did throughout the semester and reviews some of the major concepts which we learnt and then say good bye for the examinations. Let us review, what we said in the first class.

(Refer Slide Time: 02:07)



In the first class, we try to begin with a definition of what is estimation. And we said that, we want to, estimation is a kind of pheno, approximate calculation of of what; of estimation in the in, I refer to a dictionary and said that, if said that estimation is mentioned in the dictionary, as approximate calculation or judgment based on available evidence.

So, here when we talk about estimation; we are tracking, we are trying to calculate what, we are trying to calculate, signals or we are trying to calculate systems, that is why the course is called estimation of signals and systems. So, and we want to compute our, we want to we want to compute approximations of the signals. So, we are trying to leave out certain things, so what are we trying to leave out?

We are trying to leave out the noise, in in most cases. And if we talk about judgment, what judgment are we making? We are making, we are trying to generate from from information which is given to us in the form of measurements; we are trying to judge and generate several other types of information.

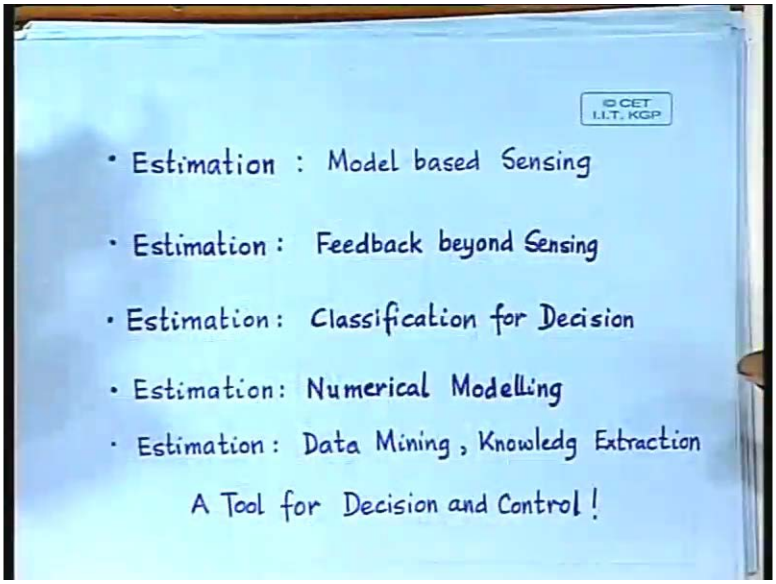
In the sense that, we are trying to predict, we are trying to find information; what we can say about, what is very likely to happen in the future or we are trying to detect, we are trying to see that, whether in some signal we are suppose, we are given a signal which is which which is mixed with another signal, whether we can detect, whether this whether a signal that we are looking for is present. We are trying to classify signals, in the sense that we are trying to see, whether certain properties of the signal are present or not.

And all this, we are achieving through computation. We are we are also sometimes, you know sometimes we cannot measure the exact signal which we are looking for; so we are trying to see whether whether through computation, we can get back the get back the original signal. So, we in other words, we are trying to compute some desirable information based on available evidence; which may contain various unknown, uncertain, sometimes undesirable elements which we typically refer to as noise.

And why do we estimate signals and systems? We estimate signals because sometimes, measurement is not feasible and or it is noisy or we want or the measurement is simply not available, because we are trying to talk about the future. We estimate systems, which are typically dynamic and complex. And all these, as we shall as we have seen in the course that, all this we are doing, using models.

So, so this is what we said when we open the course about estimation. At the end of the course, we can probably say much more many more things. So, I tried to look at estimation from various ways.

(Refer Slide Time: 05:36)



First of all remember that, what are we to what are we essentially trying to do in estimation? Essentially, we are trying to get, compute values of signals; so in that sense, it is like a sensor. What are the sensor do? Let us say, let us take a thermocouple or an RTD; if you if you put it in the process, it gives you the value of a quantity called the temperature, right. So, such devices we call sensor.

Now, here also we are we are we are trying to sense various quantities; for example, from from we have seen that, from radar measurements, we can I mean from radar measurements of position, we are trying to estimate, auxiliary. So, in that sense in that context, the Kalman filter becomes becomes a sensor for auxiliary, right. So, so we are not having any physical sensor here, we are having different types of physical sensors but we are trying to essentially sense variables.

How we are sensing the variable? By using, in this case kinematic models, right. So, that is why in a in a sense for a large number of applications; estimation is nothing but model based sensing. Why do we want to sense? Why do we want to sense something, in many cases; obviously when we are trying to sense, we are trying to get some information from the system. Why do we want that information, because because we want to take action, right.

So, obviously we do not take action open loop or you have been determine fashion; but in many situations we take action which is based on the current behavior of the system. So, so we so to be able to generate, meaningful action, we need feedback from the system that we are trying to control or we to we are trying to operate, right.

So in that sense, estimation gives us feedback but it gives feedback which is much beyond conventional sensing. Somewhere as we have seen for example, we have suppose we are trying to we are trying to probably; you have seen the case of, tool condition monitoring case. That is we we wanted, to find out, we wanted to sense to what extent a particular machine tool has worn out during its operation. Why do we want to know that, because we want to decide when we have to change the tool, so that, our performance of manufacturing is not sacrificed.

Now for that we need feedback from the tool and we need this feedback while the tool is working. So, we have seen that, to get this feedback; we have to, we can we can employ estimation, and in that sense the I mean without having any conventional, we are sensor, we are able to give a feedback for for the for this a supervisory action to the operator, by which we can decide when to change the tool, right.

So, in that sense it estimation lets you give feedback much beyond conventional sensing can. Estimation is also, for a large number of cases; estimation is actually a classification problem for for decision. For example, we have seen the case of for detection diagnosis, what essentially are we trying to do? We are looking at the signals and using estimation techniques, we are essentially trying to classify, the dynamic behavior of the system into one of two classes, either normal or faulty, right.

So, perhaps if we can get this decision, we can we can we can take very fruitful action in the sense that; we can if we find that the process is has developed a fault, perhaps we have to take a different course of a action, perhaps we have to shut down the plant, we have to we have to to do many things. So to be able to do all that, first of all looking at the input output behavior we have to classify, that behavior as whether it is normal or or whether is faulty. And estimation becomes a tool, for that classification problem.

This these commons, typically refer to signal estimation. If we come to system system estimation; we find that we are trying to estimate models and we are trying to estimate models. Typically model estimation as we have seen, involves two things. Firstly, it is model structure estimation or deciding on a model structure, and secondly trying to decide the parameters of the model within that structure.

Now, we have we have also seen, we have also made commons; that there are model is always an approximation of the system. So it is not always, so the so the actual physical mechanisms which are at play; could be could be much more complex than what is really required. What you what is required in many cases is that, we are able to predict or characterize, the system signal values using some computational method. So, for that we need a model which is which is which is able to, let us compute that the system behavior within the operating mode that we are likely to encounter.

So, for that we can without even going to the process physics in many cases, we can we can we can simply adopt some mathematical model structure and try to find parameters within that model structure; such that the input output behavior can be explained by that model structure, we we really need not bother, what are the physical, chemical, phenomenon associated with the process.

So, in that sense we are trying to build, we are trying to adopt, we are trying to adopt a at numerical approach to the model. It is true that, I mean sometimes the process physics and chemistry gives us important guidelines to the selection of the model structure; that is correct but while it is so, it we we we need not take we need not take, a records to physics and chemistry to be able to know about the structure.

For example, there are I mean, one of the one of the very popular approaches, of signal modeling is using the neural network. And it is typically used in cases, where the where the process physics and chemistry are really very complex, sometimes unknown. And therefore, we we either cannot or we do not want to really look into all that physics and chemistry. So, what we say is that let us take a neural network which is a very flexible structure, which can which can really describe or capture a lot of kinds of behavior; and then try to find out that what are the parameters in that in that neural network, so that my particular behavior can be faithfully modeled.

So, what we are essentially doing is, numerical modeling and the and the big advantage that it can give us is that, it can it can lead to a lot of simplicity modeling and it can also let us it can also let us talk about system behavior without knowing the process physics and chemistry are in; and in various situations where the where the where modeling based on physical conditions would would indeed have been impossible.

Like in the case of, again coming back like in the case of tool condition monitoring; in the case of tool condition monitoring, the how the tool actually wears through its interaction with the job is actually, an extremely complicated procedure and they are they are they are they are does not exist any well accepted method for for actually modeling that. And there are some there are some empirical models structures which again, I have been obtained basically through basically through estimation.

So, we could see that even without knowing, anything about the tool job interaction procedure and the mechanics associated in it; we have been able to model it simply, usubg is a neural network. So so so my estimation, gives us a gives us a totally alternative approach to modeling which I have called, here as numerical modeling. Similarly, estimation lets us in the in a same weighing; I mean estimation lets us discover patterns.

You know, so using estimation; if you are given a large volume of data, various variables, their their time values, we can we can we can create a mathematical object which can very compactly represent. Which can these these I mean, let us say kilo bytes and I mean, thousands and thousands of a different values; we can very compactly encode into a numerical model structure, which will be which will be good enough for our purpose, in the sense that they will be able to explain each other's behavior, using estimation. So in that sense, it is data mining; where we are trying to discover patterns, dynamic patterns, complex dynamic patterns from large volumes of data.

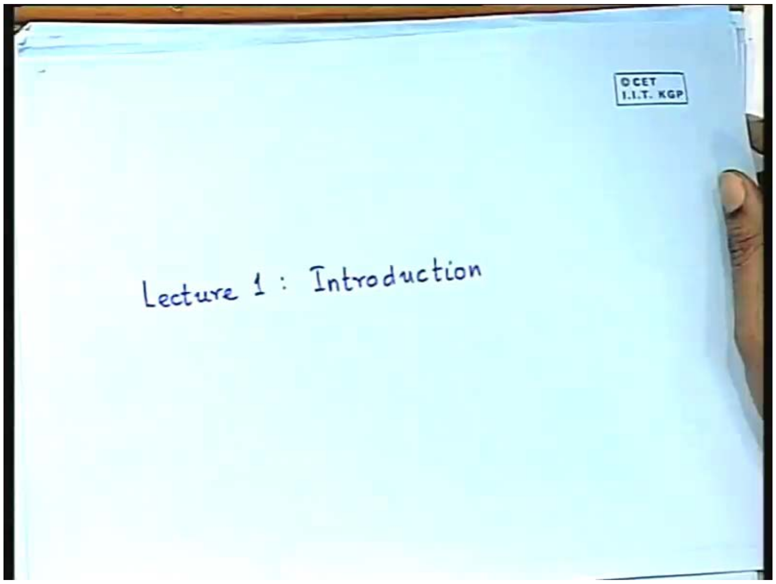
We are also doing knowledge extract. So as you can see, as we can probably reflect on at the end of the course that; this is what the techniques that we have learnt, throughout the semester, this is what it lets us do, it lets us it it basically provides a very elegant tool into the, measuring tool, knowledge gathering tool, simply from observed behavior.

And and and and that knowledge and that information; we can use for a for a host of purposes of of decision and control, that is why this course is so important. And you will find that; I mean especially now with the with computers available everywhere, I mean computers being available in in all machines, the role of information in optimal operation and control is is increasing and the use of estimation technology, in in all these applications is is becoming I mean, all most all pervade.

So, this is what we have learnt, I thought that before going to review the course; I mean how we had actually made our journey, I I wanted to emphasize emphasize, that point that throughout this course this is essentially we have learnt. We have learnt to design and implement a tool which will give us much more condensed, useful and high level information which is available in the in the measurements and which we can use for optimal operation. So, so that is what we have learnt during this semester.

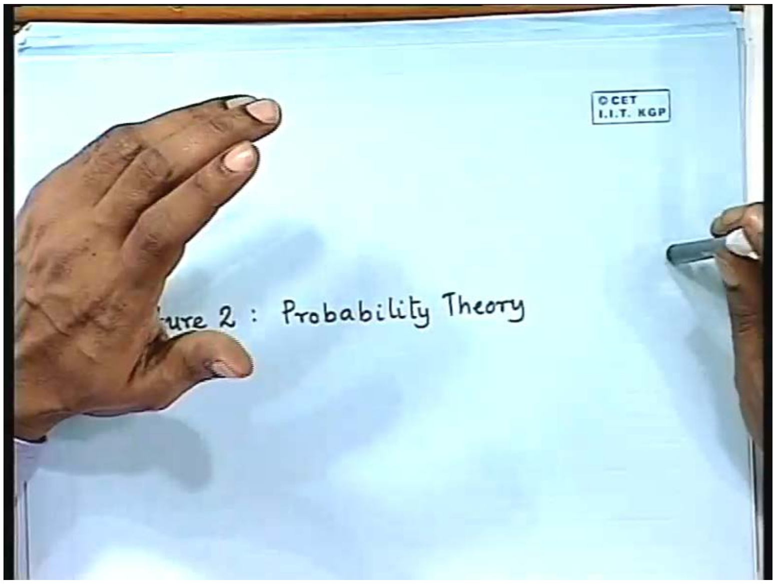
Now, let us take a review. Let us just slowly or rather quickly, go through the once just recapitulate the journey that, we are under taken in the course.

(Refer Slide Time: 18:14)



So, we be we began the course in July, the first lecture was an introduction; which actually introduce, the introduce the estimation problem and try to try to describe, what in essencesities, that is what we have just done again.

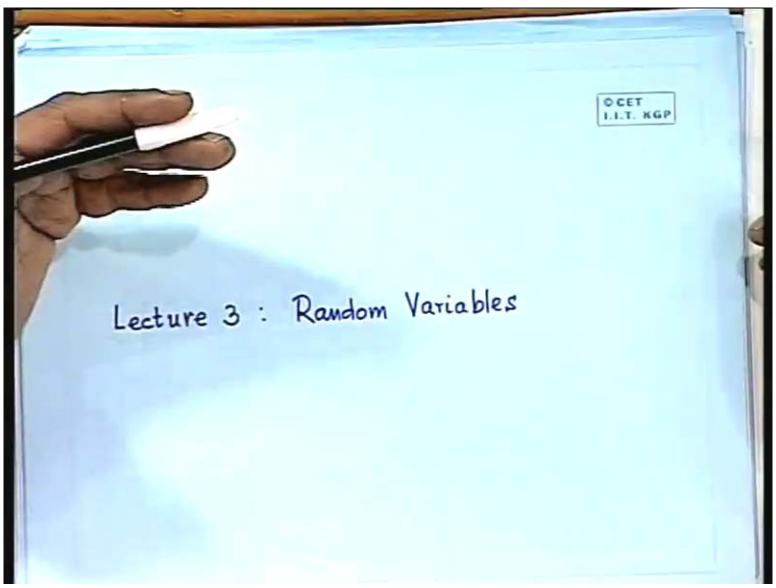
(Refer Slide Time: 18:34)



And then we began right at the beginning that is we began at probability theory, because of the fact that; estimation theory typically involves, basically inform a I mean, estimations says, that you can you can you can gather information, in the face of un-certainty and knowledge, un-certainty and noise, right. And typically, probability theoretic approach is usually adopted for a for describing, these behaviors mathematically.

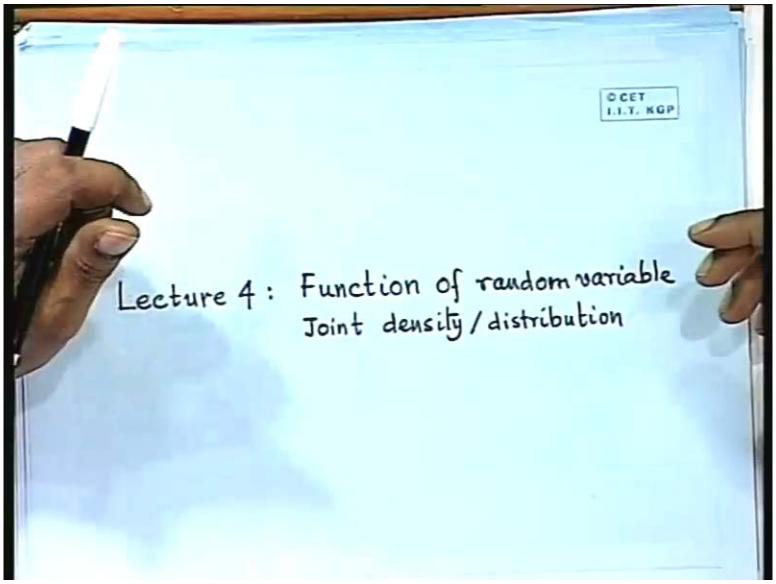
So, we without assuming any background, whatever we started right from probability theory and then went on to talk about random variables. So, these are variables which take on values and what values they will exactly take is uncertain.

(Refer Slide Time: 19:24)



So, when they are uncertain? Why are they uncertain? Because, again because there are because there are certain phenomenon, at play which influence the values that is this variables take; which we are which we are not interested in modeling either we are we are we are we are unable to model or we are not interested in modeling. So therefore, we say that these variables are random. So, we cannot exactly say what values they will take. So, so we try to characterize random variables and then from random variables, obviously; whenever we are we are we are talking about, processing any signal or processing a system we have to we have to I mean, we have to work on them, either using using various kinds of operators, right.

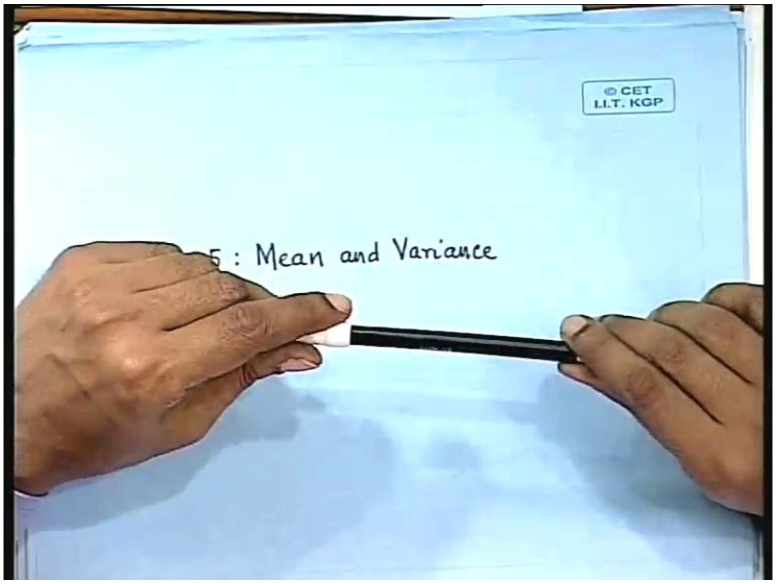
(Refer Slide Time: 20:24)



So, first of all we took a we took a static operator, in the sense that, we took functions of random variables and saw that if a random variable behaves in a certain way then how do these functions behave? What are the behaviors of the functions of them? We also saw that, if if there are a number of variables and each one of them behaves certain ways, then how how does I mean how to characterize the random behavior of the of these collection of variables, in terms of their individual behavior.

So, we so we talked about, joint density and distribution of a number of random variables. Now, it turns out that, we are always you know everywhere; it it is very convenient, if we can even if we lose some accuracy, it is very convenient to you know compress knowledge and I mean, express some rough behavior in terms of you know, one or two numbers. So, so in that sense while these variables can are are actually uncertain and they can take a values, over certain range, it is very convenient to defined one or two numbers which will roughly explain their their values and behaviors. So, so from that sense, mean and variance are very very important quantities which which which describe the the which describe two things.

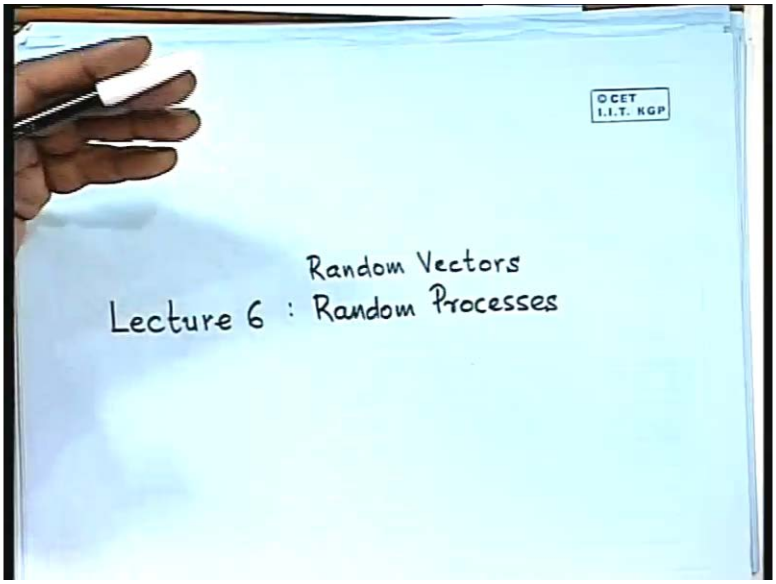
(Refer Slide Time: 21:44)



The mean is called the called a called, a measure of central tendency or that if you take ten such variables or ten such realizations of this variables, then what is an average behavior? So so in that sense, mean is important because by just one number you are able to characterize; a all the all the realizations, up to a certain accuracy. And, to what accuracy; that depends on variance which is a measure of dispersion, right.

So it says that, how much does it differ from the mean? So, we so we, define these two very important quantities which are used in the characterization of these random variables. And then we gradually, graduated towards more and more complicated objects.

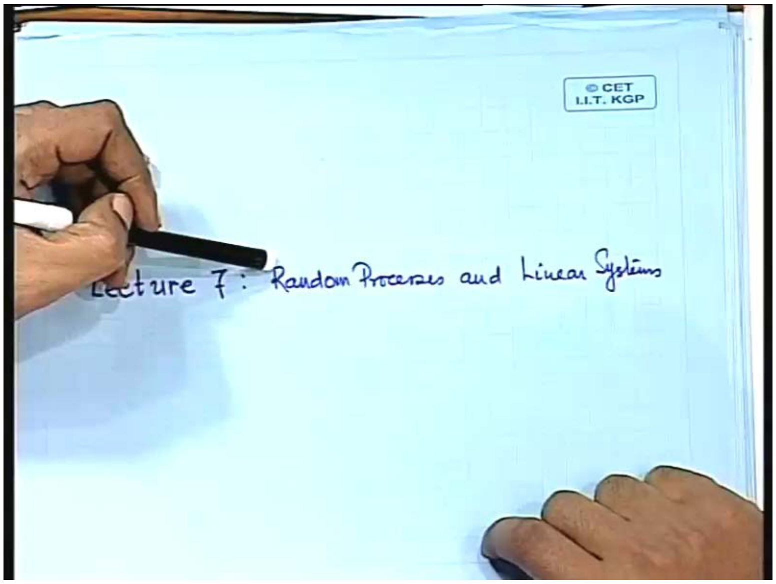
(Refer Slide Time: 22:42)



We looked at random vectors, which are you know I mean, an organized collection of random variables and then we looked at random processes. So, so you know random variables are single values, while random processes are actually random functions as we have seen. So, from from random processes; now these just like variables will be operated by functions. Similarly, functions of time will will be operated by dynamic operators.

So, we have to we have to characterize that, how these functions when they are when when they go when they go through various processes, how their time behavior gets gets transform by this processes? This is of is of, I mean a great interest to us. And essentially to keep thing simple, we we always consider linear dynamics. So, then we came to see, how this random processes behave when they are pass through linear system.

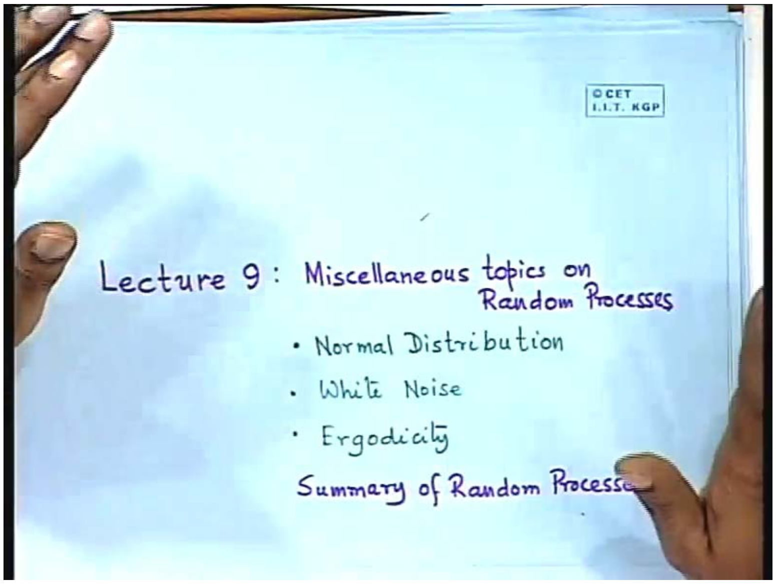
(Refer Slide Time: 23:46)



So, if you are given the power spectral density of the input, can we calculate the the power spectral density of the output? Because, that is very important because in many cases; we what we want is that, we want to characterize the characterize the process. So, we characterize the process, essentially in terms. So, we say that this is a process which which when fed with an input of this kind, will produce an output of this kind.

So to be able to characterize these, this phrase, this kind, we need to we need to characterize these random processes and the and they are they are they are processing through linear systems. So, we did that. After that, we did solved some problems; on lecture eight and finally we reviewed, some some some very important special topic on random processes, before actually going on to an estimation problem.

(Refer Slide Time: 24:53)

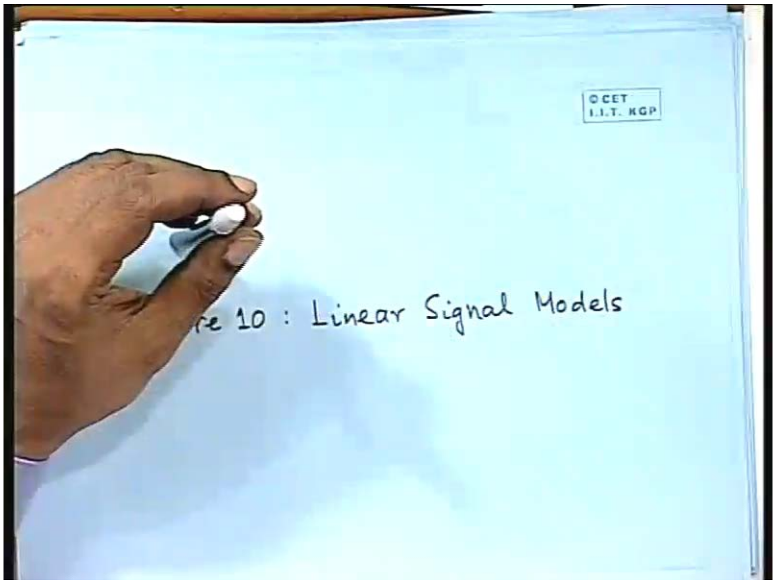


So, we talked about the normal or the Gaussian distribution; which is which which is very important because especially, because it it has certain very interesting properties when when operated through, linear systems, like retains, its Gaussian character. We also talked about, white noise because because white noise is in a sense; you know like an I mean, it is a it is a characterization which is a which is a counter for the impulse, and it leads to considerable analytical simplicity in our analysis.

We also talked about, ergodicity; which is a you know, we we have to talk about in many cases, it turns out that that to be able to characterize average behavior, it is very inconvenient to talk about ensemble averages because, usually we do not have so many realizations of a process. So, we instead, we often records to time averages of a single process and substitute that for ensemble averages.

So, now I mean essentially when we want to do that, we have to make an assumption about ergodicity. So, we so we reviewed this the concept and summarized our studies on random processes before we before we became ready for handling the estimation problems.

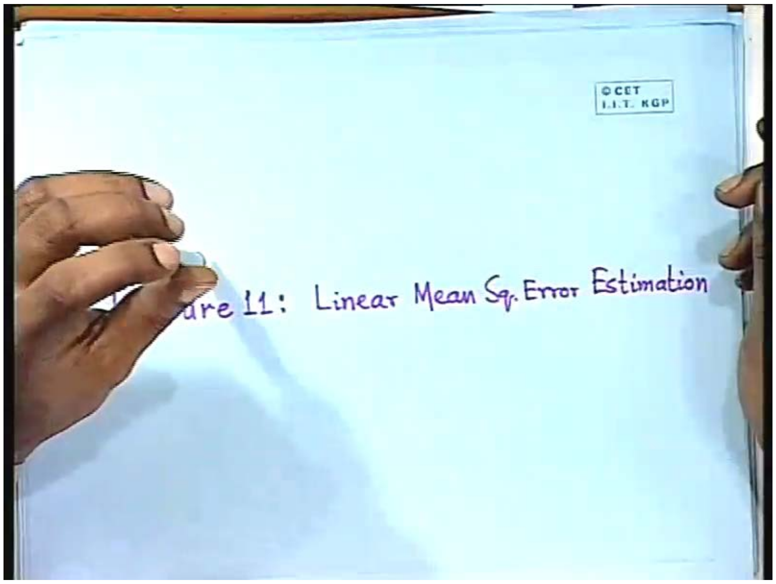
(Refer Slide Time: 26:30)



In lecture ten, we again looked at, some some further on linear signal models; especially because we found that, if we want to characterize if you want to characterize output properties with input properties, through linear signal models then there are then there are certain certain certain kinds of linear signal models which will which which are very convenient, which can very conveniently model a very wide class of, I could have put behavior.

So, we looked at that and we looked at the all zero, all pole and pole zero models. And at that point of time, we we considered our first real estimation problem which is the linear mean square estimation.

(Refer Slide Time: 27:22)



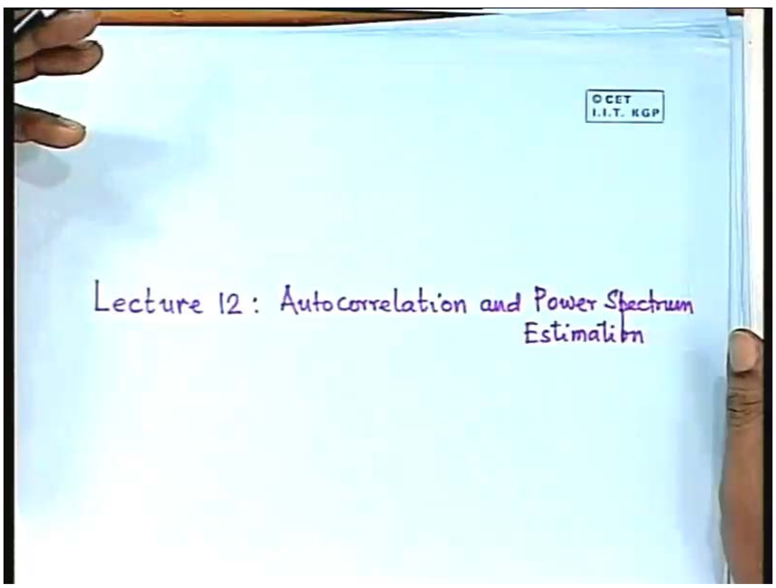
Which says that, if I am given measurements of a certain quantity, as us as values which are sequence over time, then using this using a finite sequence or an infinite sequence of these value, can I estimate another quantity, can I do that? So and if I can do that, can I do it using a linear model; and if I want to do it using a linear model, if I only restrict to linear models, what is the best I mean, how to obtain the the best estimator, which gives me the minimum square error of estimation, on an average on an average over ensembles, again.

Which means that, if we do these estimation with fifty, hundred, two hundred, different signal sequences and get corresponding error sequences, and compute square errors, and and average theml; which is the value of this estimator, what are the parameters of this estimator which is going to give a least mean square mean square error. So, that is the minimum mean square error problem, which is which is well known and which an essentially reduces to a winner filtering problem. And we obtain a solution of that, using the so called normal equation.

At this point, we you know in in many cases it happens; that we want to use frequency domain information, we want to do frequency domain modeling because of the fact that, various I mean complex, differential equation properties in the time domain, become become become algebraic so that, their manipulation becomes much user in the frequency domain. So, it turns out that; we we need to understand how to estimate, what are known as auto correlations and then take the auto correlation in the frequency domain to get the power spectrum.

So, we need needed to study, how to estimate auto correlations and power spectrums; and then obtain frequency domain models of systems, in terms of the power spectrum of inputs and the power spectrum of outputs. And it turns out that, under certain very rather mile conditions, a lot of power spectra can be actually modeled, using linear pole zero models. So with that, we studied the auto correlation the an and power spectrum estimation.

(Refer Slide Time: 30:19)



We studied two approaches; one based on smoothing, another based on ensemble averaging. And we found out also nicely, that if we are given a sequence then these two approaches are are are of equivalent; in the sense that, the they they lead to the same kind of you know, loss of frequency domain or time domain accuracies. So, now we are we are we are we are ready to explore models in the frequency domain and since we are mostly working in using in the discrete time frame work; where we have rather than having a signal, we have a sequence of sample values.

So, the frequency domain descriptions of systems in the sample domain; is the is the z transform, so we recapitulated the z transform. And we also recapitulated some a basic concept of algebra, relating Eigen values or Eigen vectors; which becomes a useful in in many kinds of analysis of stability, innovations, etcetera, that was lecture thirteen.

And in lectures fourteen, once we have understood our Eigen value, Eigen vector concepts and how they can be used, to you know orthogonalize a sequence; we introduce the concept of innovation. The concept of innovation is is very interesting, in the sense that; it it it it transforms the information which is contained in a sequence of data values into a form, where the where the the information is is not really spread on spread on each each data sample. That is each data sample brings a new information.

The the the the sum total of information, remain the same as the as the old sequence but the but now; this this this transform sequence which we sometime called call an innovation sequence, every component of that sequence or every sample of that sequence, brings a new information, so that information which is not obtainable from that of any other sequence. So, this you know kind of you know, compartmentalizes the the the information and again lets us do analysis in a in a much more deep and effective manner.

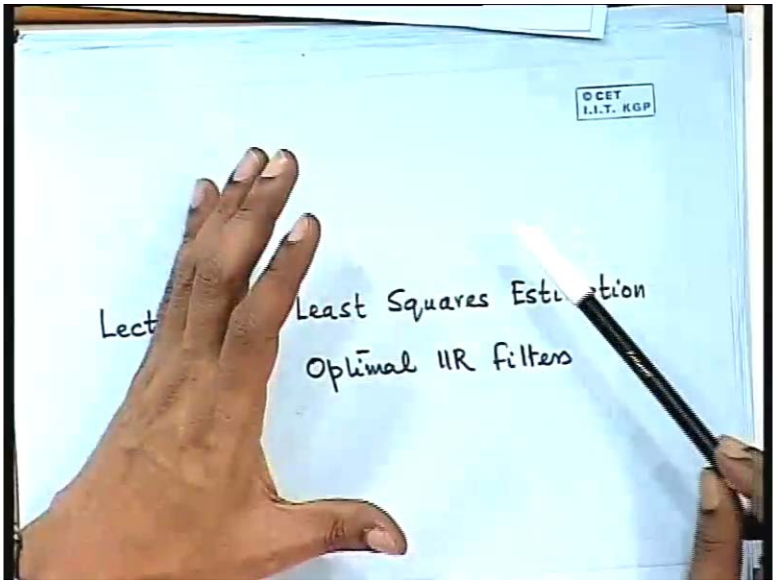
And then we saw that, how this what is this innovation, in the context of the a linear minimum mean square estimation; which we have already studied and found that the the the optimal minimum mean square, I mean the minimal minimum mean square estimation actually leads to the the error being perpendicular to the data. So in the in the sense that, it becomes orthogonalized and each of this errors actually bring us new information; which we need to incorporate while we are while we are updating our parameter coefficients.

Coming from linear mean square estimation which actually; you know linear mean square estimation is in a sense optimal over all possible data sequences, I is true because it because it minimizes the error over over over an ensemble average, that is correct. But at the same time, as you have seen it requires knowledge, it requires knowledge of the properties of the ensemble, which may or may not be available.

So, when it is not available when it is not available what are what what we can do is; we can say that okay, I do not have the data available for the whole ensemble, but at least can I can I obtain can I obtain can I obtain the the best data, which is best for this data sequence? So, if we are trying to optimize the error over the whole ensemble, we have minimum mean square estimation. If we are trying to optimize the error, only only for the given sequences; then we are doing least square estimation which are very very you know, related.

So, we saw that how far, you know we initially, we actually talked about the linear mean minimum mean square minimum mean square estimation, in the context of FIR models, all all zero models.

(Refer Slide Time: 34:31)

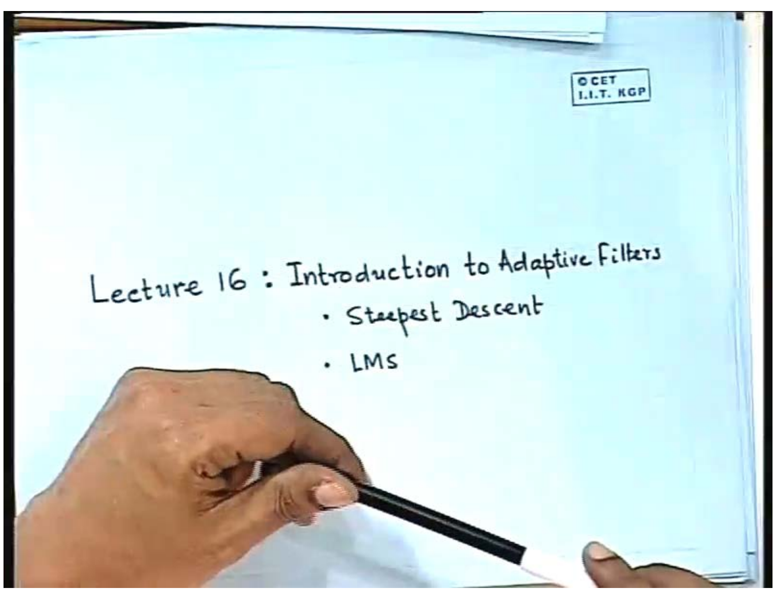


Then we saw that, how we can how we can estimate optimum optimal IIR models; and we saw that that actually employs the concept of innovations, the concept of pre-whitening and a lot of other things. Having described fixed filters, I mean how to how to estimate fix fixed filters? We went to tackle a new situation where the properties of the data may be changing over time; in which case if we employ a fixed estimator, as soon as the properties of the data change, that estimator does not remain optimal.

So in the sense, actually we cannot really compute a linear minimum mean square estimator because the because the matrix R and D; that is the cross co variance and the cross and the auto co variance matrices, they actually do not exist because the because the data is non-stationary. So, if we have non-stationary or quasi stationary data, where the where the where where these ensemble properties are slowly changing; we need to we need to also keep our filters not fixed but slowly adapting with the with the with the data properties, so to do that you need adaptive filters.

So so essentially, what you try to do as we have seen, is that you try to compute these properties these ensemble average properties; online from the data that you have available. And we looked at, two two different, we took at two different approaches.

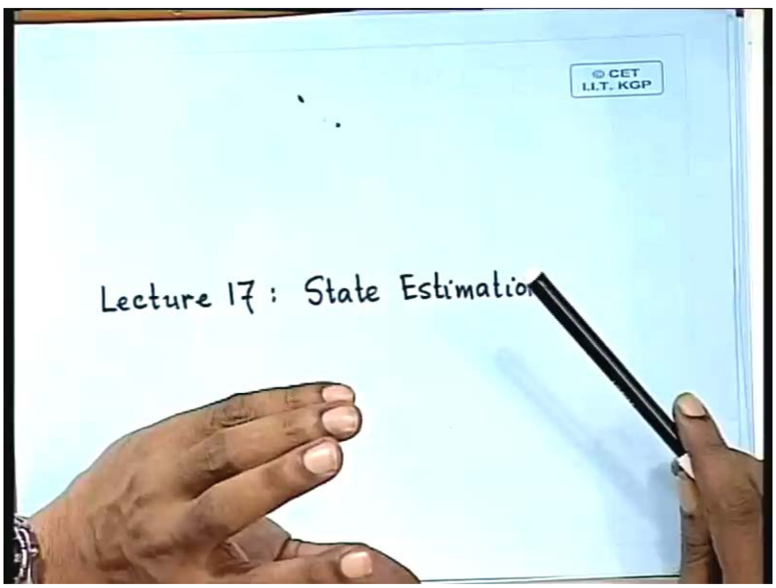
(Refer Slide Time: 36:30)



One is the steepest descent approach and the other is the least mean square algorithm. So far we have been talking about, what is known as input output model in the sense that; we a assume that the measurements are the inputs to our filter, and the quantity that we want to estimate is the a is the output, but there are lots of situations where the for the process, that we are trying to estimate we have some measurements.

Now, we want to estimate other quantities like, like a state, you know we we for example; we get again coming back to that, target tracking example which we did, we are trying we are getting measurements only a position. So in that sense, since we are getting measurements of position; position is an output but we are interested in in knowing not only an accurate value of the position, but also of the of the other states of the system, that is the acceleration and the velocity.

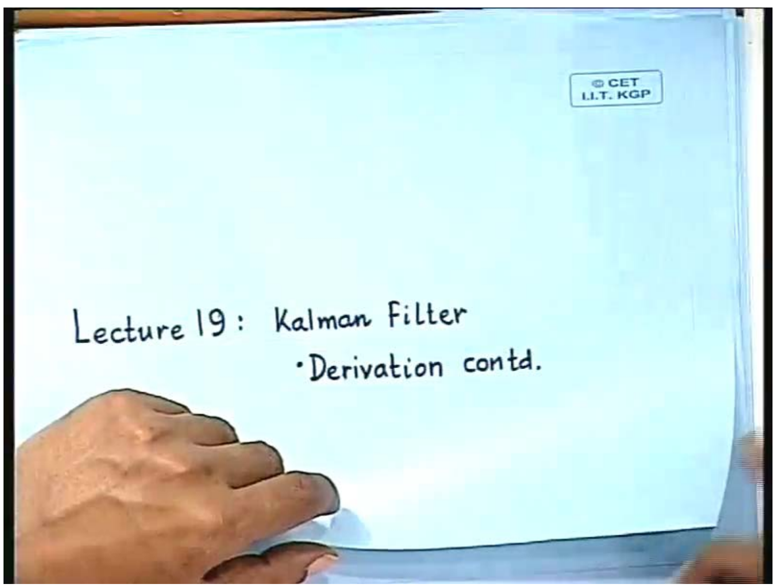
(Refer Slide Time: 37:45)



So, so from a from a set of measurements which are actually; a in a combined expression of the states, we want to get the states. This is a this this a very common problem and a and a very very important problem for a for a large class of dynamic filtering cases. And we we took that problem up and and, after describing the basic rules of state estimation in a in a in a in a deterministic frame work; we look at the Kalman filter which is an optimal linear estimator under certain assumptions, for a stochastic version of the state estimation problem. That is when we have when we have uncertainty in the process input, in the sense that we do not really know what what what the process inputs are; so we assume that the process is being excited by some random signal and we also have uncertainties in in the measurements.

So, in the presence of these uncertainties; how to obtain an optimal estimate of the system state is the problem of Kalman filter. This is the very very well used and well known algorithm. So, we looked at it in great detail and derived the Kalman filter over two classes, lecture eighteen, nineteen.

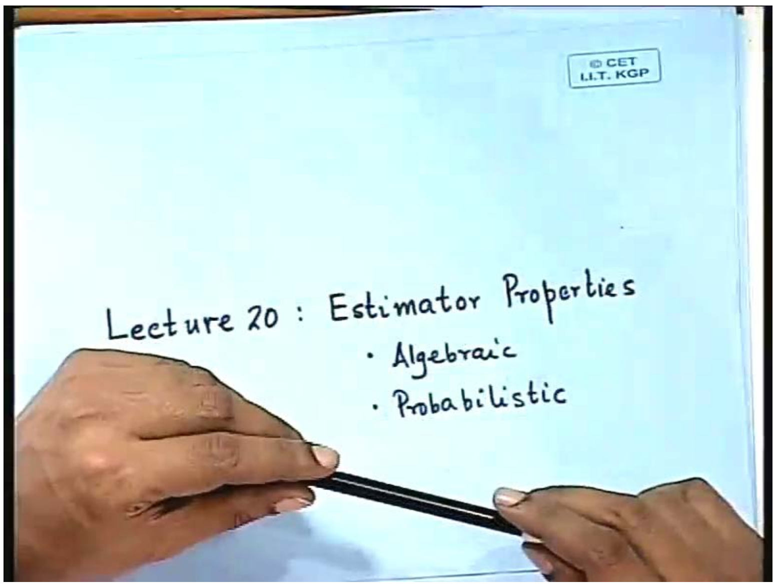
(Refer Slide Time: 39:14)



And finally derive the Kalman filter equations, from from from from the basics, from the basic least square optimization view point. We adopted that view point; because there are act, I mean actually as it turns out that; the I have been mentioned in the class also, that the Kalman filter can be derived in various ways and a various assumptions.

So, we found that a the least square, the least square optimization view point is; you know one of the easiest to understand because, it does not assume any make, does not make too much assumptions, about the noise statistics and the same time does not require, a lot of noise about probabilistic descriptions of signals. Having derived the Kalman filter; we looked at its properties namely, its namely two properties. First property was that, again some algebraic orthogonalitic properties, that is the the again the estimation error is orthogonal to the measurement sequence, as well as to the estimate sequence.

(Refer Slide Time: 40:16)



So, so in that sense it is we have be able to in that sense, once once when the the the estimation error is actually orthogonal; to the state estimate sequence then it turns out that, we cannot extract any further information from the measurement sequence to to improve our estimates. So, in that sense, the the estimate becomes optimal. So, we so we looked at that property and we also looked at, a very nice you know very nice probabilistic property; which says that under certain Gaussianness assumptions of the noise, the Kalman filter is actually a conditional mean, that is given the measurements, that estimate is the best best estimate.

So, we also looked at this probabilistic property, we estimated. Now, if you put that all first of all there are, we this is the this is the basic Kalman filter property; which have which has many assumptions in it, and so so necessarily we have to take a look at how we can you know relax of the assumptions because in a in a in a in a in a real situation, if you want to employ the Kalman filter, we I mean our our our contest may not really satisfy, all these all these assumptions that the Kalman filter needs.

So, in that case we need to really make a little little modifications in the filter and we looked at two cases; one is that the is that, the measurement noise is is coloured and second that the process noise and the measurement noise are correlate. So, we saw I mean, very simple methods by which these two cases can be tackled and we also took took a look at; what happens to the Kalman filter in the study state or when does it become a become a time invariant filter because that that is very interesting.

Because, mainly because of the fact that; we are we are we are first we are in many cases interested to understands steady state properties, properties which hold for long periods of time. And also because once we have a time invariant filter, it becomes so easy to analyze control loops. If you have a if you have a Kalman filter as as as I have told just now, that an estimator provides feedback, for what for for control or for guidance in aerospace.

So, so so here basically, it constitutes the feedback transfer function. When you when you a feedback, you actually feedback, choose through some transfer function, which is typically refer to as etchers. So, the so the estimator stands in that box which is the etchers. Now, if it is the time invariant filter, then it is so much easier to to analyse the behavior of the whole control loop; and so it is interesting to know, under what conditions the Kalman filter becomes time invariant.

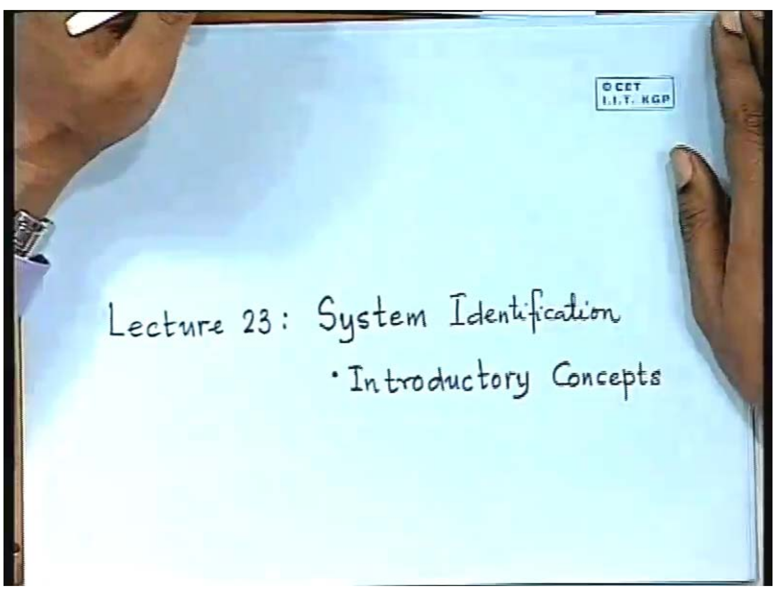
So, we looked at that and finally we came to some applications. We in detail, we looked at a we looked at a target tracking case and we found out; indeed that the that the that the Kalman filter can indeed remove a lot of noise, in fact what was very impressive in our class, to see that if we did not have the Kalman filter and if we had a noisy process, I mean noisy position measurement as we as we had and we try to differentiate that to get velocity of an acceleration, it would have been completely hopeless.

So, that shows how cleverly the Kalman filter algorithm has been designed because it gives us a pretty good estimate of the a velocity of an acceleration; and it has been widely used in in in many many applications, one of the most important applications been being the aerospace systems.

We looked at the various computational issues. A issues where issues where numerical problems can actually degrade performance. We looked at the square root filter. And we also looked at other variance of the Kalman filter; like the like the extended Kalman filter, which is useful for non-linear systems or for simultaneous straighten parameter estimation where where the parameters of the system at are simultaneously changing, and we want to we want to estimate the parameters as we estimate the states.

At this point, we so we have seen two important classes of signal estimation; one based on input output models and the other based on streets based models.

(Refer Slide Time: 45:18)



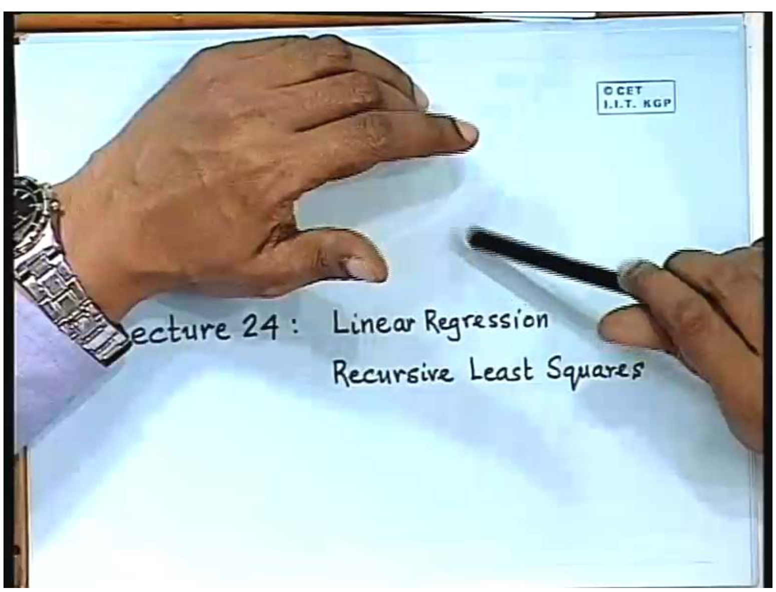
At this point, we we we concentrated towards the towards system system estimation, which is the which is this which is the second major component of work port. So, we first describe some

some some some basic introductory concept; like like like you know realizing that, the that the that the model is after all and and and and approximation of the system, which only remains valid under, in a certain operating condition, certain kinds of the input that you give and for certain purpose.

So and we also describe the setting of the of the of the system identification problem that is you have you have set of input, output data; you have sometimes, you have a model structure given and you have you have a model structure given and you also have a performance criterion given. And you want to find out, the the best model within which fits that model structure, that is called a model set. And we we we want to find out, the best parameters for the for the model within that set.

So, again we we we we we began with the simplest possible model structure; that is called linear regression models, I mean actually a lot of lot of systems actually, fit this model structure and it has been widely widely used in in various kinds of systems and control.

(Refer Slide Time: 46:39)



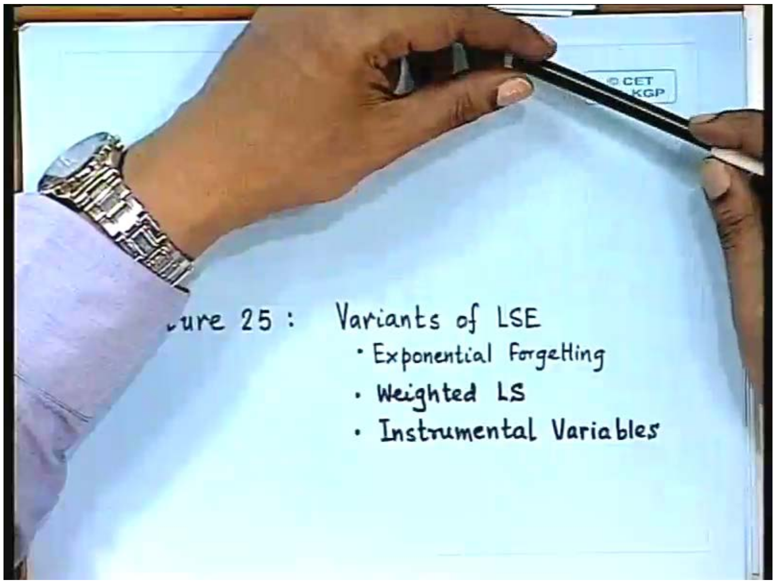
And then we looked at a very simple and well used algorithm, called the recursive least squares. Now, the recursive least square algorithm, works in the base case but there are number of cases; where we would like to, well say in a several variance of the algorithm can really improve performance.

So, we consider some some variations, in the in the in the basic least square algorithm in the sense; that we consider exponential, for getting exponential for getting, lets us lets us you know keep the it rather than finding a model which is good good, equally good over the whole time risen. If you feel that the that the system is slowly time variant; then it make sense to say that, the parameter estimates which have been obtained now, should best explain the data in the recent past and need not explain, the old data to that a to to that good an extent.

So, in that can case, we are we are what we are essentially saying is that; we I I need to put more emphasis on the signals which are of recent past, and need to explain them better and can perhaps, ignore the distance past signals to to be able to make a compromise. So, that is achieved by exponential for getting. Sometimes, we would require that some some parameters, suppose we are interested in; suppose we are interested in the D C gains, so we say that we want to we want to estimate, the low frequency properties of the system very accurately and we can perhaps sacrifice the high frequency properties to an extent.

In which case, you have to have you know we you have to have weighted variations of the least square. Then one of the one one of the problems of the basic least square algorithm is that, it gives what is known as, parameter bias.

(Refer Slide Time: 48:41)



In the sense that even, if the even if the model structure is the same as that of the system, because the noise is actually correlated with a data; you are if you if you if you simulate the data using, even even if you simulate the data in a in a computer and then use the exact same model structure, you will not get back the parameters which is used for simulation.

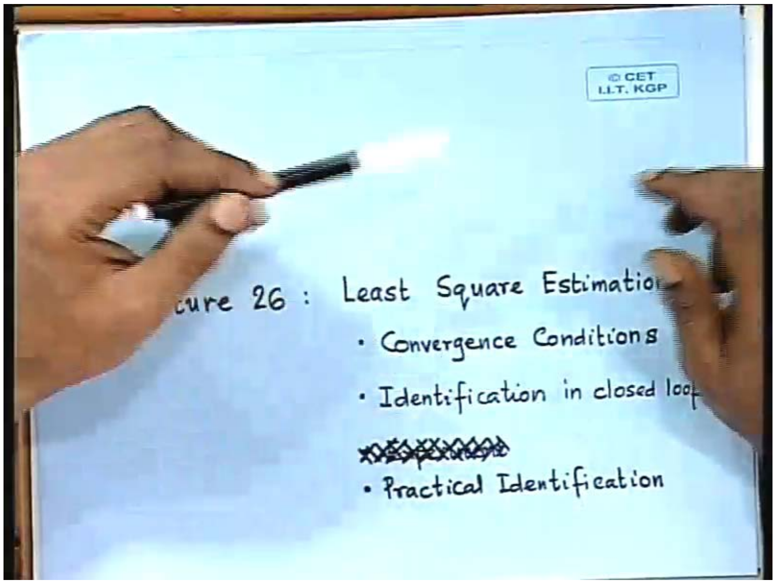
So, so you have to get a parameter bias and this essentially occurs because the noise is gets correlated with the measurement. So, there are there are there are again various ways of you know, tackling that. One one approach is to is to is to model the noise, which is you know very similar to the case of; the case of the correlated measurement noise case, we that we saw in the Kalman filter. In the sense that, we have to then once once the once you find that the noise is actually correlated; you have to model that that how this correlated noise was actually generated.

So, you have to model the noise generating mechanism and you know a enhance the, I mean argument your model structure with this noise model and then go on estimating both the system parameters, system model parameters and the noise model parameters. There is there is another very elegant approach to to remove bias and that is the method of instrumental variables; where you actually use a signal, I mean in in certain places rather than using the measurement vector, you use a vector; which is which you call the instrumental variable vector and which I mean, it has the property that it is very well correlated with a measurement, but it is highly uncorrelated with the noise.

So, so we are actually taking two approaches; in one case we are trying to in including the noise model, we are trying to whiten the noise. In the other case, using we we rather than using the actual data or the actual measurements which are correlated with a noise, we are using a slightly different data, so that it remains, it becomes uncorrelated with a noise. I mean, the advantage is that you do not have to use a noise model but then the disadvantage is that, you have to you have to again compute, this this you know so called, instrumental variables.

So you know you I mean, if you want to get something you always lose something; it is very difficult and it usually does not happen, that to be able to I mean you get something without sacrificing anything else. Then we looked at, you know some some special topics on least square estimation. Number one, is the is the condition of convergence that is; we when is it guaranteed that the least square algorithm will converge?

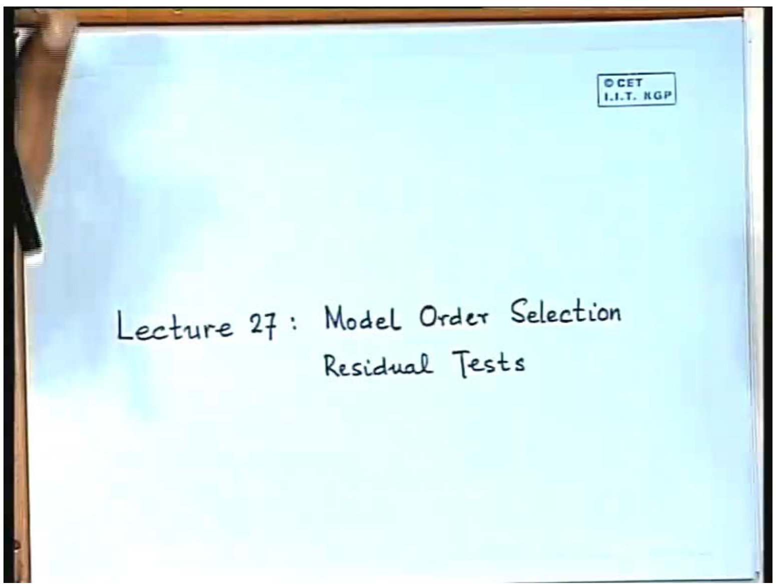
(Refer Slide Time: 51:37)



We looked at, how we can identify a system in close loop; because it it because it may not be always possible because there there are certain problems are occurring, close loop in the sense that the typically in a close loop the noise tends to get correlated with the measurements because the because the inputs have been generated from the outputs, so the output noise fits into the inputs.

And there are in certain cases, if you are very, if you use very simple controllers then you can get into what is known as the loss of identifiability; if you use close loop identification. And then we looked at various for you know practical identifications strategies.

(Refer Slide Time: 52:18)

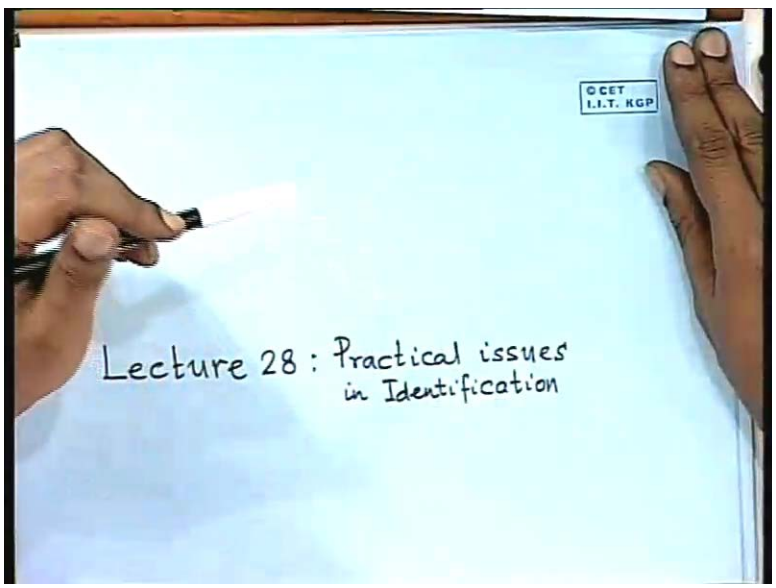


Finally, we looked at model order selection and then that is how do we know, what models, I mean; that the least square algorithm typically assumes a model structure. So, we saw how we can select model structures, essentially, iteratively. And also by looking at certain structural algebraic properties of certain data matrices; like you know correlation matrices and I mean, examining their ranks. We also can look at the residual because we know that; if our parameter estimates are going to be good and under the model structure, then the then the residuals we have certain properties of whiteness.

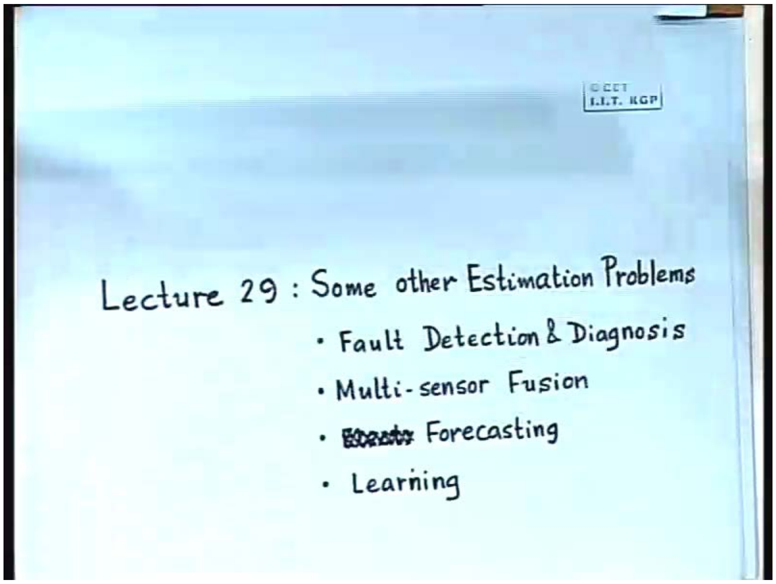
So, if if after long estimations that whiteness does not come then we can perhaps; infer that there is something wrong with the model structure itself, go back and change the model structure and then again go through the go through the estimation process and again change the residuals. So again it is an iterative process.

Finally, we saw that how suppose, we are actually given a situation where we have we have a practical system and we have to you know connect wires and we have to get the data; it may be using data activation cards and actually write a code and do system identification. What are the various steps, what are the various issues that we have to go through, before we can get a model which will be good enough for our purpose? So, we looked at the practical issues in identification.

(Refer Slide Time: 53:38)



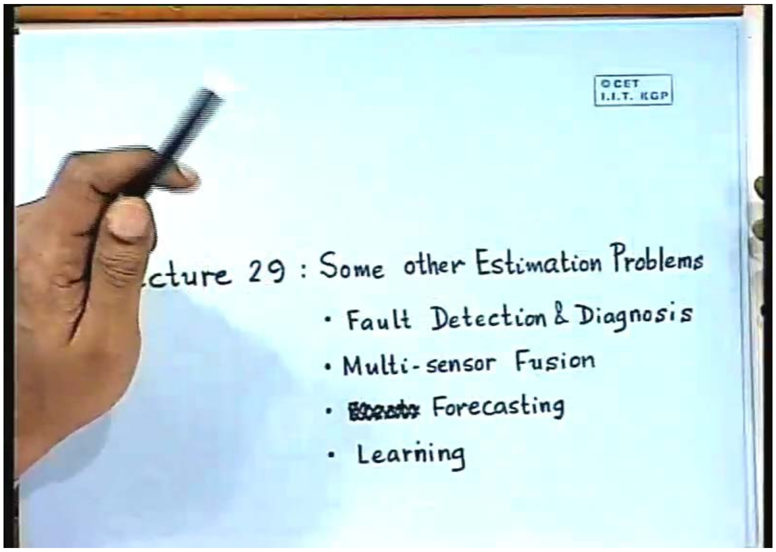
(Refer Slide Time: 53:44)



So, this you know kind of completed, what we wanted to complete in our course and as I have told you always; that it is it is very important to know in a course, I mean especially towards the end of the course; as to what we have not learnt because it because it it will always turn out, that you have any course in in any area, you cannot learn everything in that area.

So, it is so it is good to take a look at, very brief look at many other things; which we did not get time to look at you know in detail. So, in the last lecture we looked at some other some other estimation problems which we did not really seen depth.

(Refer Slide Time: 54:17)



For example, we looked at the fault detection and diagnosis problem; which is very important especially, for some safety, critical and vision critical applications. We looked at multi sensor fusion, and here here it is that we looked at the tool conditions monitory; example that we repeated where you know, using more than one sensor and using clever estimation, we can we can really construct, what we can say, what we can call a super sensor, for which we cannot have any physical device. So, we can enhance the the the individual properties of each sensor, just by combining them and then using an estimator.

We can use forecasting, forecasting lets us you know, look at the future and probably get prepared for it; take take take preparatory actions, very much used in various kinds of industries, stock markets, electric utilities. And then look at higher, I mean a kind of estimation which is really a higher than signal system estimation which is learning.

So and that brings us to today’s lecture. So, this is what we did in the course; essentially I think that with what you have been told, you are very well equipped to go ahead and tackle any estimation problem that may arise, in a given engineering contest. So, with that I would like to close the course, it was a very pleasant experience for me. Thank you very much.