**NPTEL**

**NATIONAL PROGRAMME ON TECHNOLOGY ENHANCED LEARNING**

**IIT BOMBAY**

**CDEEP**
**IIT BOMBAY**

**ADVANCE PROCESS CONTROL**

**Prof. Sachin Patwardhan**

**Department of Chemical Engineering,**
**IIT Bombay**

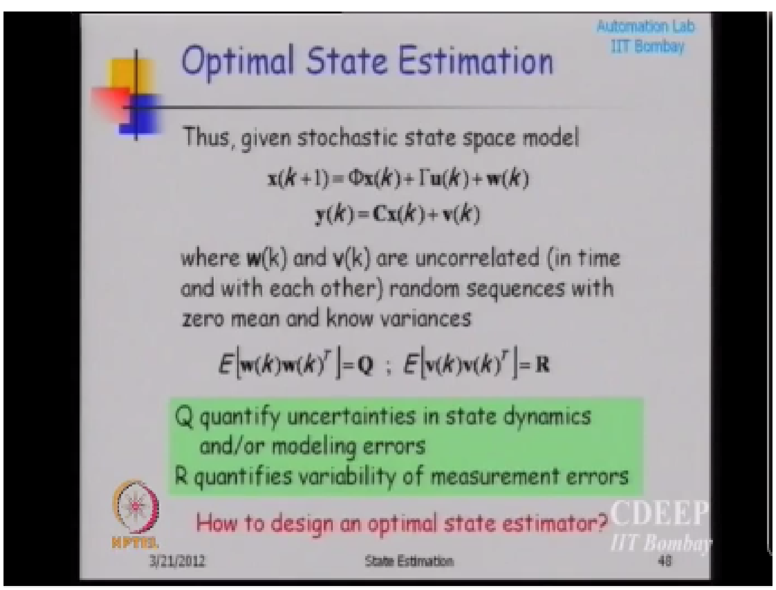
**Lecture No – 19**

**Soft Sensing and State Estimation**

**Sub Topics**
**Kalman Filtering**

So we were looking at this optimist estimation and we arrive at this stochastic state space model.

(Refer Slide Time: 00:23)




Here wk and vk are stochastic process means white noise process, classic example of system where xk will be a correlated stochastic process, why it is correlated? Because xk is correlated with xk – 1, xk – 1 is correlated with xk – 2 by this difference equation xk – 2 is correlated with xk – 3, so there is the correlation between xk xk – 1, xk xk – 2, xk xk – 3 okay. So even though wk and vk are white noise process xk is not, here we are assuming that we know characteristics of wk and vk.

We know that wk is assume to be a 0 mean assume white noise process vk is 0 mean white noise process, we know the co variance of wk and vk okay, so we know this. So this is my model okay, given this model I want to come up with optimal estimates of states x okay, in a such a way that I use the information about these two noise sequences, I have this information that co variance of wk I do not know what is exact value of wk.

See there is the difference between these two inputs one input is uk, uk is known input, wk is unknown input, we do not have any measurement yet we have characterization of the input in terms of a noise and in terms of mean and variance. So we have some model from it, it is not that we are completely blind about what is happening in terms of unknown inputs okay.

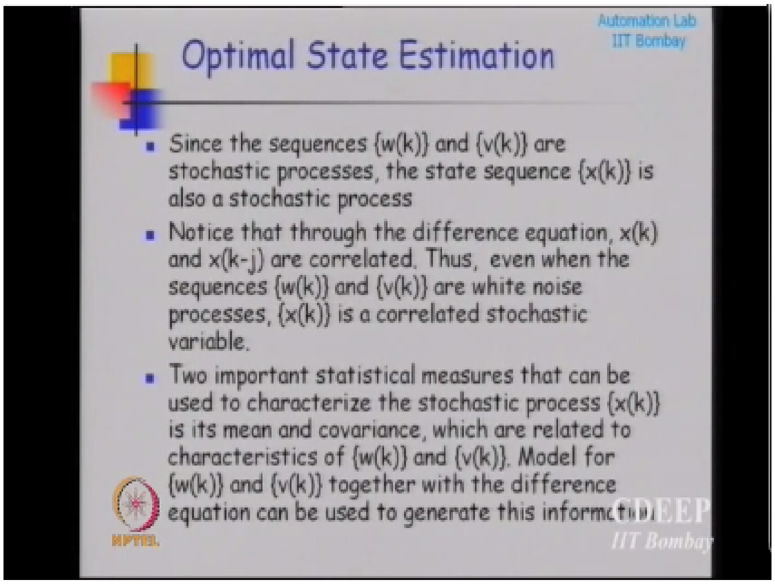
So now how do I systematically incorporate the information of this covariance while doing state estimation how do I optimally estimate the states okay. My task is following okay, I want to somehow when I do state estimation okay, my measurements are correlated or corrupted with noise okay and there is an input that goes into x.

I do not have measurement of input okay but remember effect of wk will percolate to yk +1 actually, so effect of the state disturbance is present in yk okay and when I do state estimation I want to somehow account for the wk, because it is an input that is going to state dynamics okay. I better account for it okay at a same time vk is measurement noise it is complete dirt in my data I want to remove it okay.

So even though we call this as a noise it is customary to cal this as a noise state noise and this I the measurement noise, this vk is unwanted okay where as wk you would like to have an estimate and compensate the state estimation for wk. even though you do not have an measurement you would like to have an estimate okay.

I want to compensate my state for wk, I want to use this information Q and R while doing so, Q and R some uncertainties that represent in the input and R quantifies the uncertainties or the variability of the measurement and then question is how do I design a optimal estimator that is what we are looking at.

(Refer Slide Time: 04:21)

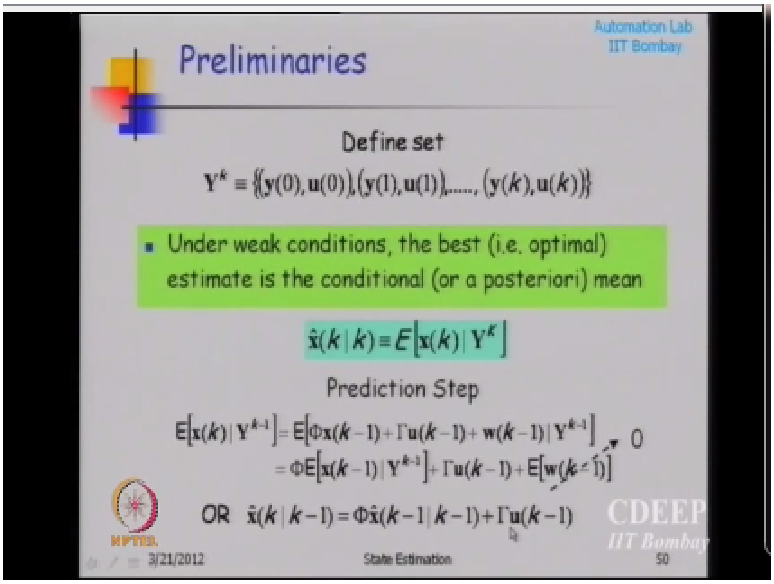



First of all wk and vk stochastic process and by the virtue of the fact that the difference equation is driven by stochastic inputs okay xk is also a stochastic process okay, xk is the function of uk and wk, wk is the stochastic process then the xk is the stochastic process there is the randomness in xk okay. Then this I just talked about that you know through difference equation xk is going to be a colored noise what is the colored noise.

Time correlated noise if you try to find out auto correlation between xk and xk – 1, xk and xk – 2 it will non 0, auto correlation function for wk auto correlation, what is auto correlation for white noise? It is equal to co variance for lap 0 and any other co relation is 0 okay, this is not going to be the case with xk is going to be correlated random variable, it is stochastic process correlated in time.

And we have to uncover it is characteristics okay, that is one of the main thing, so what I want to do is I want to somehow link the statistical characteristics of xk okay and rk and vk that is what I am going to do okay, that is what my next task.

(Refer Slide Time: 06:12)




Now before we reach the final point okay I have to do a lot of preliminaries and you have to bear with for some time for the entire picture to become clear okay. So I am creating some intermediate results keeping them aside and then I will combine everything into you know this development okay, so first of all I am defining this set, this set is y0, u0, y1, u1 this set of measurements that have collected over time.

I have this data with me okay I am going call this set as $Y^k$ this is not y:k this is just notation okay, it only means that set of all data collected from time 0 to time k. So if I write time k + 1 which means data collected from time 0 to time k + 1, if I write data collected from time 0 to time k – 1 okay that is the notation. Now I what I want to do is well when you are doing an observer you are going to use measurements to correct the states.

I am going to use the measurements information to correct the state estimation not correct the states, the true states cannot be corrected the estimated of state is going to corrected using feedback from the measurements okay. So what you can show is that this estimate of x will be function of both wk and vk, why it is function of wk, because wk is affecting the state dynamics why it is the function of vk because we have the feedback correction coming in okay.

In the observer I have the feedback correction which is based on the measurement okay and just imagine why entire data is required? Why I am saying that the entire data is required because when I estimate y x 1 I am going to use information of y 0 and y1. When I am going to estimate x2, I am going to use y0, y1 y2, when I estimate you know x10 I am going to use information of y0 to y9 y10.

So measurements are going to be used to reconstruct the state estimates okay, so what I want to find out is that estimate is a stochastic process, estimate itself is the stochastic process. Differentiate between the two things the true system dynamics is also stochastic process estimate of x is also another stochastic process, there are two different things okay. Now we are saying that the best estimate that you can get okay.

Is the conditional mean that means generate conditional density of x, condition of measurement that you have obtain okay. How to generate this, it looks very abstract, how to generate this condition density? I will work out the algebra so that we will be doing but if I can get this conditional density and if I take it is mean that is the best estimate of the xk, so somehow I want to arrive at this condition mean okay.

So I have two steps I have prediction step and I have a correction step we are going to develop a filter, now what I am doing here is just I am using this difference equation just go back here, I am going to use this difference equation and find out conditional means okay. so this is the difference equation now written at time point xk, so forget about this yk – 1 for the time being just look at this.

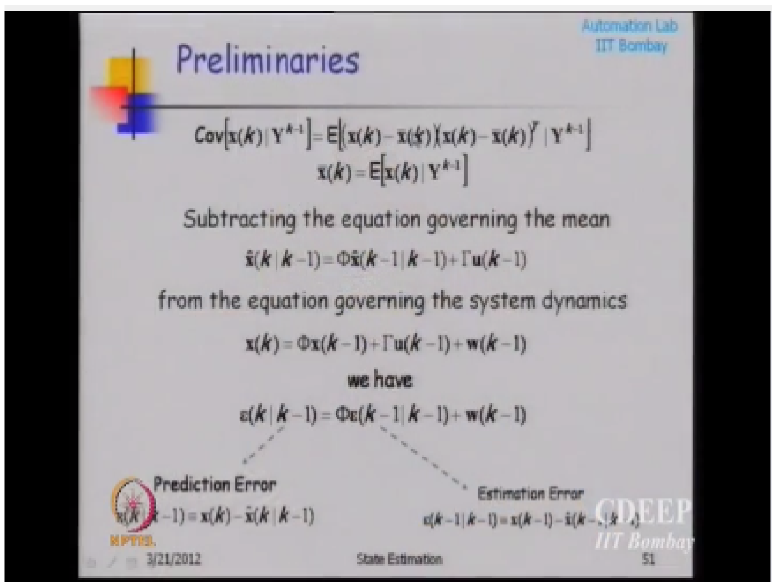
Xk okay = φxk – 1 + γ uk – 1 + wk – 1 this is what we have okay, now if I use information upto yk – 1 okay what is the meaning of this $y^{k-1}$ all measurements upto k – 1 have been used okay. Conditional density of this so I am going to take an expectation operator on the right hand side, okay so look here φ si the matrix so I have taken it outside and I am writing expectation of xk – 1, yk – 1 is everyone with me on this okay.

This is the deterministic input we know what is uk, we know γ , so this comes out okay, so expectation of this is nothing but this itself, it is not stochastic variable and what about this guy expectation of wk is 0 mean wide noise what is the best value 0 okay. So this is 0 so which means I get a recurrence relationship, what is the recurrence relationship? The new mean, the new conditional mean, see this is estimate of x.

Mean of xk condition on yk – 1 what is the meaning of this, mean of xk conditioned on yk – 1, see here yk – 1 is appearing okay. because we used here set upto k -1 this is conditional mean of x upto using information upto k – 1 if φ times conditional mean exact k – 1, k -1 + u input that is gone in okay. So what I am saying here, see do you remember sometime back when we talked about stochastic process we said that for a general stochastic process the mean can be function of time.

Mean is time varying okay so this xk is stochastic process who is and yeah last slide because there si no information of wk – 1 yk – 1 because wk-1 will effect xk, xk is in future yes. So w k -1 it is effectively present in yk but not in yk – 1, so understand because that one is 0.

(Refer Slide Time: 14:17)




See to characterize a stochastic process I need two measures very important mean and variance okay what does, forget about this is the multi variant time varying process, just one signal random variable what does mean tell you, most probable value right, what does variance tells you? it tells you roughly physically, just do not forget that physical interpretation. Larger the co variance okay the possibility of x being anywhere in that band.

You get a larger band around the mean where x can take the value, smaller the co variance or the smaller the variance okay; x will be closer to the mean. The value of x which will actually occur see expected value of x and actually value x will take different. X is the random variable when you say that mean is the best estimate, they are estimating predicting okay. So you know that is what given the density function that is what the best you can say.

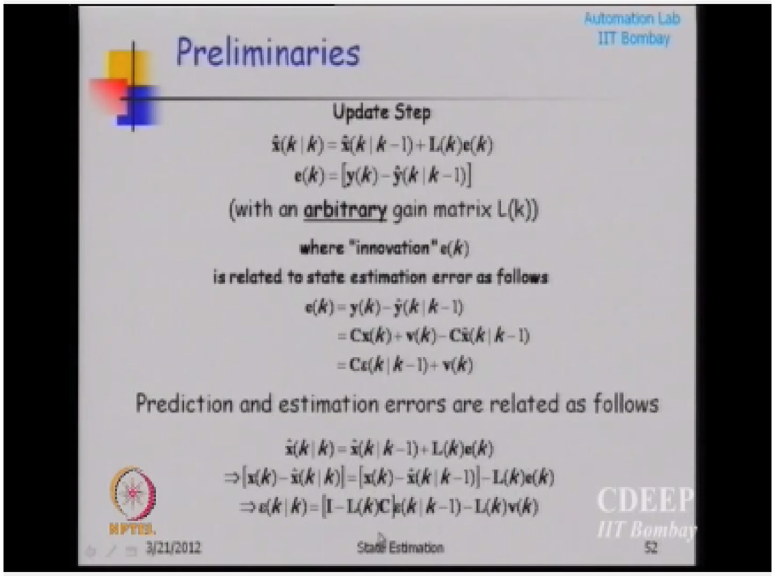
And co variance will tell you the spread so do not forget this particular idea okay, though we are going to get into more complex math okay. So now what I am going to do is I want to find out now co variance, I found out mean and now I want to find out the co variance okay. so to find out the co variance I have to subtract the mean okay and then, so this is my mean this is our mean changes as a function of time.

This is the original process I am going to subtract this from this okay if I subtract this from this, see what will happen if I subtract this equation from this equation, this γuk – 1, will cancel these are constants which are adding same constants been added to the both the equations, so this will cancel w k – 1 will remain okay. So if I actually find out this difference okay, I am defining this difference x k -1.

So this is I am going to call it as a prediction error okay, I have predicted value of x right now I have not done any correction okay and this is estimation error, is xk true value of x – estimate of x, condition or estimate of x okay. Using information upto k -1 so I am getting the difference equation that relates new error with the past error and wk. so this error itself is a stochastic process, now we have to find out you know.

You know if I write this here you know it goes outside the slide, I have practical difficulty, if I write here this – expectation of this inside this x bar is nothing but xk given yk – 1 okay, so mean I want to write mean okay, conditional mean but at conditional mean if I write inside I have trouble okay. I am just keeping that result and moving to the update step.

(Refer Slide Time: 18:40)




We will sort out how to come to the co variance, what is going to update step? I am going to take the predicted estimate, I am going to merge it with measurement using a game matrix lk. How do I chose the game matrix lk that is my design problem. I want to choose lk in some best possible ways okay, so that the estimated is optimal.

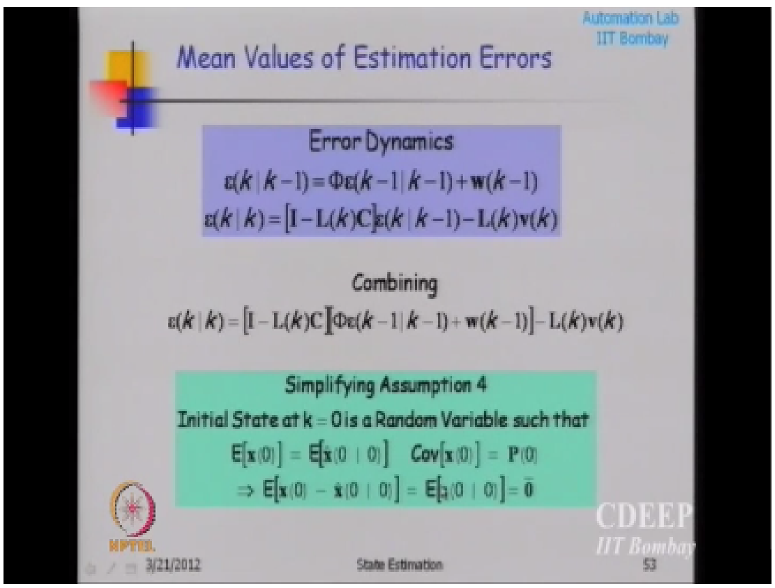
Now what is this optimal that what we will see, is everyone with me on this, now I am going to fuse y which is measurement what is this y hat k given k – 1, estimate of y using information upto k – 1, how do you get this estimate, c x x k -1. So now I am writing this equation right now for some orbiter L. my task is to decide L itself where what is ek is typically call as innovation.

Ek is nothing but yk – yk k – 1 okay, what is yk, yk is cxk – 1 okay, so this is cxk – cx hat k -1 which is nothing but c ε, k given k – 1 + vk okay. What you can appreciate here is that these estimates now are going to be functions of wk and vk. Why it is going to be the function of wk? Wk effect is present in yk is not, wk is a some noise entering you know the real plant, so it is effect is present in yk okay.

Wk – 1 effect will be present in w effect what I mean to say is disturbance effect will be present in yk past disturbances effectively present in yk okay. so now what I have done here is I have written this x hat a given k is k – 1+ LK ek but ek is nothing but c εk + vk, so I have derived a combine expression this is just an algebra okay if you have notes just look at it is just an algebra I have just derive a combined expression which is combining this step to okay with the.

So now what do you have at the end of this, at the end of this we have two difference equations just look here what we had in the last equation this equation was relating error k – 1, k-1 to k k -1 okay, in the next equation. I have relationship between k k -1 to k, k so I have found out difference equations between successive errors okay how the errors are govern by is this okay everyone with me on this yeah okay.

(Refer Slide Time: 23:00)




So now I want to find out mean value of the error what is the mean value of the error? so my error dynamics together I have written these two equations together. I have these two equations these two are coupled equations you see this k – 1 gives me k k -1, k k will be used in at time k + 1 k and that will give you k + 1 and so on okay.

So everything depends upon ε0, if you know ε0 okay you will know ε 1, 0 if you know ε 1, 0. If you know ε 1, 1 you will know ε 2 1 and you can start rolling. So I have just combined this and I am writing now k, just substitute this and this okay, so now the equation will be k as a function k – 1. Now here is the simplifying assumption okay, what I am saying here is that expected value of x 0, see when you this is the tricky thing okay.

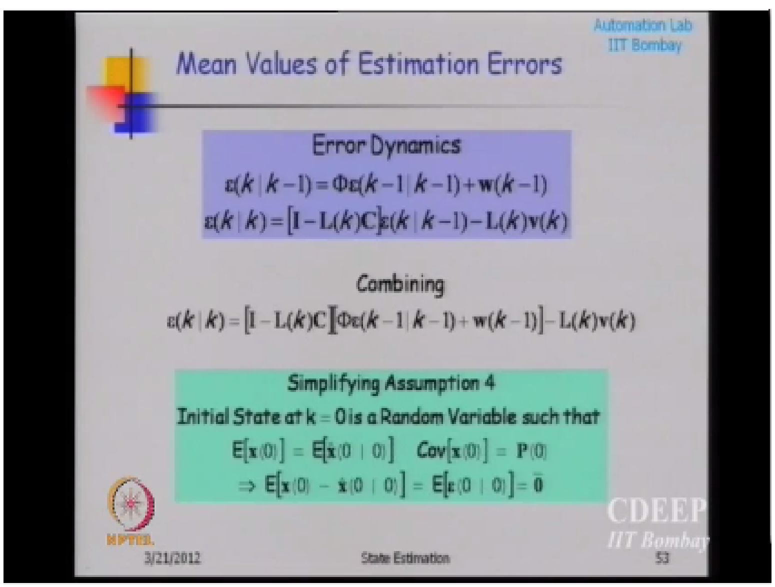
Expected value of that true value at time 0, see I have to give an initial guess right to start the difference equation, how do I give the initial guess. So when I give a guess I am it is the random variable I am guessing okay. I have to create a random variable okay, I am making an assumption that the guess comes from the same distribution as the two distributions, and they have the same mean okay.

So x0 is the random variable initial guess that I give is random variable but both of them as same mean okay what does it mean and then the co variance of you know are equal. So I am just saying that you know the initial guess is the random variable whose mean is same as the true mean, whose variance is also same as the true variance, this is my simplifying assumption. So but see the just assume this for the time being and see beauty of this algorithm.

I mean this is something like this algorithm as you can say change the world okay, so and see main thing in you know in science or in mathematics is to make right assumptions okay. if you are able to make right signifying assumptions you know because when you model there is always a disconnect between the reality and the model.

So you should know where to draw the boundary, if you know that boundary then you is there okay, so I am assuming see that the expected value of the initial error is 0 okay. Now what are the consequences of this assumption?

(Refer Slide Time: 27:27)



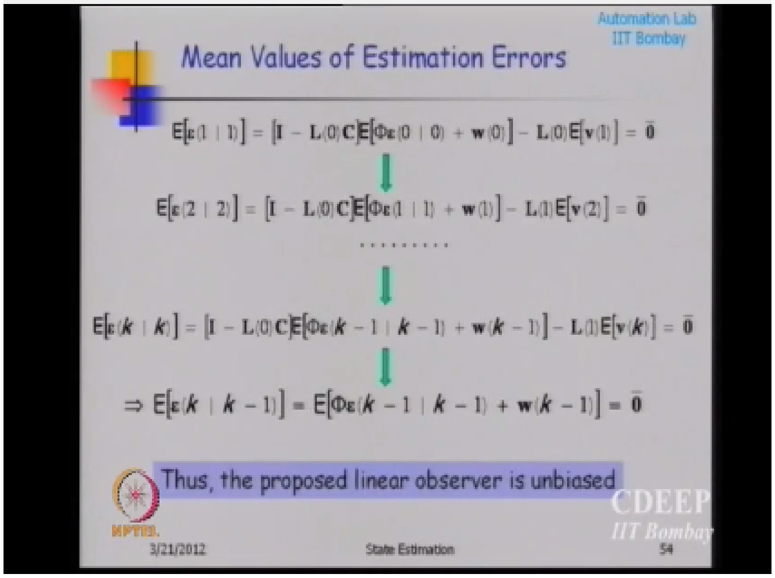

Just see what is the consequences of this assumption, what is e 11, see I am going to use this difference equation now together with this assumption what is the assumption? Expected value of the initial error is 0. I am not saying the initial error is 0 I am saying expected value of the initial error mean of the initial error so I have guessed and error that you committed, if all of you guess and if take mean of those guesses.

Why x0 is the stochastic variable? See x is the stochastic process see time 0 is t=some arbitrary time right let say it has started at some time - ∞, see we are talking some arbitrary 0 time and then starting okay. See you are in an imaginary world in which the stochastic process has started at some time - ∞ when the universe started.

So now you decided to start to do the calculations, so 0 is arbitrary fix today but something is happened in the past right, let us take an example you know x is the temperature in room and k is you know that is a day k – 1is yesterday and k + 1 is tomorrow k + 2 is day after tomorrow okay. So I decide something to calculate using today, so today is my 0 okay. X0 that is the true state today is actually the function of all the randomness in past.

So what is x now is x is the stochastic process where I place 0 is my choice okay something as been happening before time 0, al that information is contained in x0 okay, I decide to use equation from today, so I will start using it x k + 1 tomorrow, but the past history is there in x0 right and it is the random variable okay, we are assuming that we have some information about it through p0 we know in some sense we are saying that we know it is mean value. There are different ways to interpret these equation different books you find different equations,

(Refer Slide Time: 30:36)




So I am going to start using this equations rolling in time, so now what is the expected value of e 11, will be I – L 0 here x expected value of φ x e 00 okay + w 0, what is the expected value of e00 our assumption is that is 0, it is the white noise okay. So what is e 22, see e 22 will depend upon e 1, 1 but e 1, 1 is 0 okay, so what you can show is that, if you make an assumption okay that initial error had 0 mean.

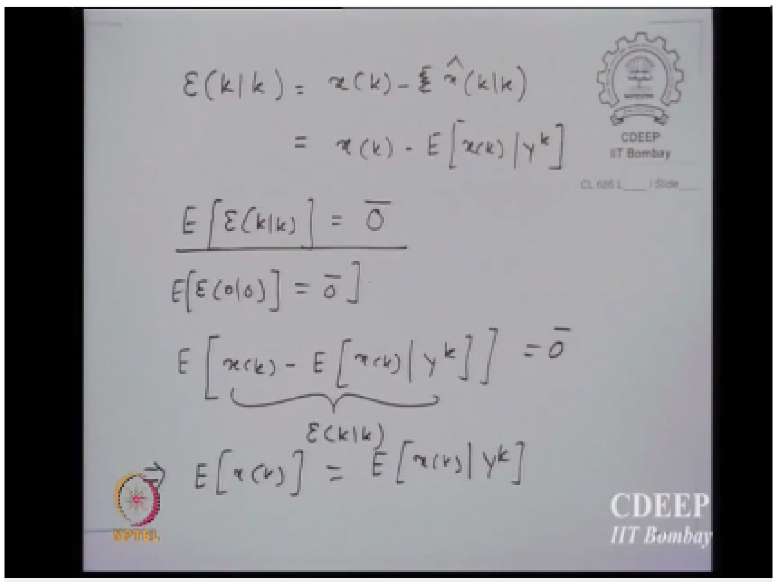
If you make an assumption that initial error at 0 be then all subsequent error also has 0 mean okay so this is called as the unbiased estimator okay, expected value e00 is not equal to 0, expected value e00 most likely, no I am not saying that, I am saying that there is going to be an initial error but that error is random error whose you know distribution as characteristics that if mean is 0.

So actual value of the error at times 0 will be down from the distribution actual value is different from the expected value. That is the assumption okay way to think about it is that suppose I tell all of you see, I tell you that actually x0 comes from the Gaussian distribution okay and there is the method to draw a sample from the Gaussian sample from distribution okay. If you ask for a sample from Gaussian distribution of certain means certain co variance it will give you a sample okay.

I tell each one of you to generate one sample okay and is subtracted it from the true value, I know the true value and I have given you the distribution you generate the sample okay, then I will find out the mean of each one of them and take your sample, his sample her sample I will take the difference between true value which I know and the samples which you have given me okay. Then I got all possible e 0, 0 see I asked you guess the temperature in this room.

I know the true value you can guess, each one of you will give a guess and then I will find an error, now you have given me a guess and if I take mean of all possible error values between the true which I know and the guess which you have given that mean is 0, that is what I am saying, is that clear okay is everyone with me on this upto here okay. So this error it is mean value is 0, now, so I found out the mean of the error what the meaning of the mean of the error is is 0?

(Refer Slide Time: 35:02)




So I am saying εk k is xk – x hat given k right this is what our definition is, that is xk – expected value of xk conditioned on yk right. So I am saying that expected value of ε k/k = 0, that is why I proved under the assumption I made, how did I prove this. I could prove this using difference equation together with the assumption that expected value of e 0, 0 = 0, this together with the difference equation follows right is that okay.

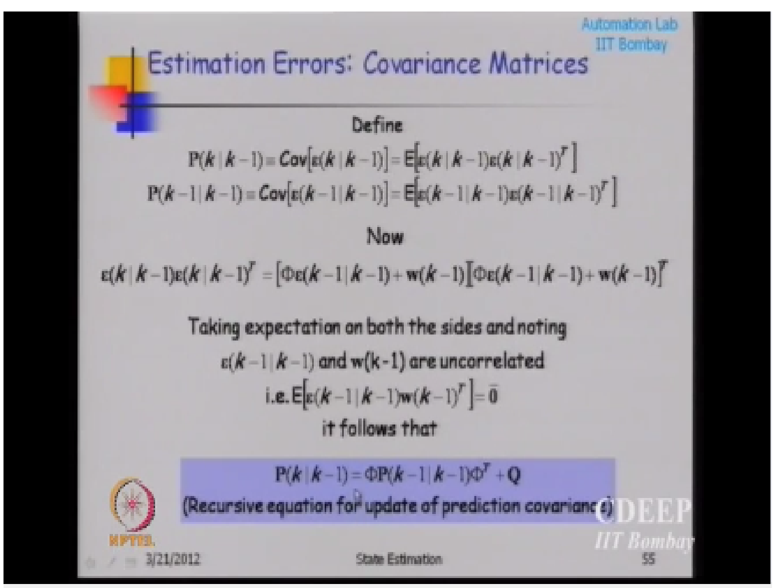
Now I am just going to interpret this so expected value of xk – given yk = 0 right 0 vector this is nothing but εk right okay or this also implies that expected value of xk = okay. so conditional mean is going to be a unbiased estimate of x. there is no some constant coming up expected value of xk is same as conditional expectation of xk, condition on the measurements, expected value of xk is the random variable.

Conditional expectation of xk is also another random variable I am showing an equivalence what I would say this that do not expect that you will understand everything in one short okay, so things will sleep in slowly okay. So at some point you have to accept and proceed and then but what is expectation already is same value, expectation is the deterministic quantity. Expectation once you found an expectation deterministic value.

Just think about it I am just saying that estimate is the random variable truth random variable both as same means, the way I am constructing estimate the mean the estimate see what this say is the conditional expectation of x is same true expectation of x, so it is fine to find out Conditional expectation of x. if I want the mean value of x I could as well find the Conditional expectation of x.

I am just showing the equivalence between the two okay, say it again I am not following what you are saying, let us discuss it later, and see if it becomes clear, it does not become clear you ask again okay. So now to find out covariance what you do?

(Refer Slide Time: 39:56)



I am going to find out what is the expected value of this quantity is 0 okay expectation of ek is 0 right, ek is 0, ek / k-1 is 0 also both of them are 0 because if 1 is 0 the other is 0 because they are relate to algebra. See we showed that expected value of ek k is 0 and by this correlation expected k / k -1 is 0 both the error of expected value is 0 okay. So now I am going to use this to find out co variance,, so I am defining this co variance matrix, see how co variance matrix is defined?

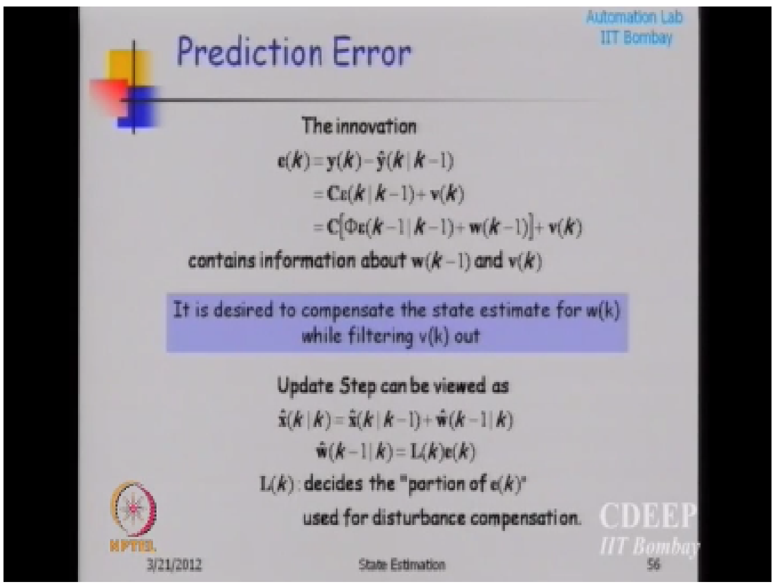
Co variance matrix is defined as expectation of k k -1 k -1 εk why mean is not appearing here? Because there are 0 mean both the errors are 0 mean okay. So I just have to take εk k -1 x ε k -1 and here okay. Now I am going to derive a recurrence relationship between the two okay, this is the stochastic process I am going to derive a recurrence relationship. So I just use the difference equation k – 1, k- 1 is same as φ k – 1 x wk this whole thing into T.

I am just multiplying left hand side multiplying right hand side okay and then I am going to take an expectation, now just think about this wk -1 and error k – 1 k -1 are not correlated just go back and look at that is why I want to have an equation I cannot keep going back in the slides, if you have the print out just look at the equations. These two w k- 1 and φ k -1 , k-1 they are not correlated.
So expected value of this quantity is 0 and from this it follows you get this recurrence relationship okay very simple derivation you just do multiplication of these quantities these are vector and matrix quantity so you have to very careful when you take transpose and all that okay, if you take transpose and expectation you will get this recurrence relationship that is predicted co variance okay is 5 times updated covariance at k -1 okay + covariance of wk.

Why is this Q term coming here? Because you will have wk transpose okay see whatever comes you will get when you do this multiplication just go back here, you will get φ then you will get term between ε x w you will get two terms between ε x w and then you will get one term between w w transpose. So there are 4 terms out of which 2 terms cancel only 2 terms remain.

So which 2 terms cancel by the virtue of the fact that wk – 1 and error at k -1 are uncorrelated we get this recurrence relationship. Pk -1 is prediction error, so this is co variance of the prediction error, this is co variance of estimation error okay this is the recurrence relationship okay. Previous mean is corelate to the new mean same way I am finding the recurrence relationship between the covariance. So I have characterized process in terms mean and co variances okay I now I have to do one more step okay.

(Refer Slide Time: 45:02)




Now look here what is ek? Ejk is yk – innovation okay so this is ε k -1 + vk okay, what is the mean value of ek? You take the expectation of ek what will you get? You will get 0, why will you get 0? See because I have shown that the error yk – yk – 1 is the function of wk – 1 that is what you want to know, it is also function of vk. See I have been just doing the substitution ε k -1 look at your slides.

Now if I take an expectation of ek it will be c φ x expectation of this quantity is 0, expectation of wk -1 is 0 and expectation of vk is 0 okay. So actually what I want to do is, first thing I want to convey here is this ek it contains information about the disturbances but it also contain vk, so we now want to separate out information about vk and wk.

Somehow try to compensate the estimate of the state using information for wk – 1 I want to estimate wk -1 using ek and that is what actually you are doing when you are putting this lk times when you are doing this you are actually constructing the estimating of wk and you are actually correcting it here, it is an interpretation okay, why I am doing this correction? See this is the correction spell.

This correction is the ek okay, ek contains the information of wk – 1 okay but it also contains vk, so I want to multiply it by a factor what is the role of this factor? This factor will help you to correct for wk – 1 it will try to filter out vk. This is the interpretation which I am giving for what we are doing. Now how do I decide Lk is the golden question, we have not answer that yet that is what is going to be my next task okay.

So what is the expected value of ek is 0 what is the can you find out co variance just do it, find out co variance vk tell me what is the find out co variance vk? There should be some little break in between, which one so it is estimate of noise at the previous instant using information upto current instant. So there are 3 things in the estimation okay.

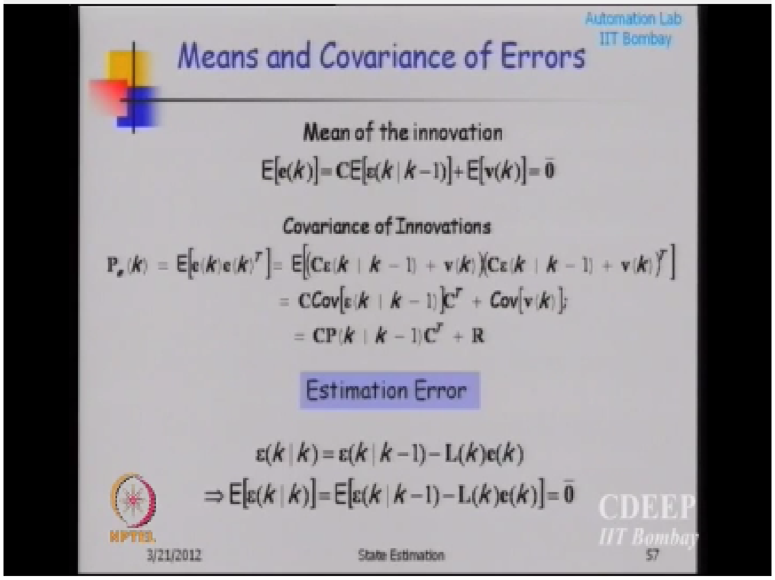
One is called as prediction estimate / k -1 is the prediction estimate k / k is the current estimate or filtered estimate and you know when you are estimating something in the past using information upto future that is called as smooth estimate okay and see I have collected it is like saying you know you collect the data of temperature in this room starting from say 1$^{st}$ of the January till to date you have time series you collect the data.

Now you have the model and I ask you to do prediction of what is an estimate of temperature tomorrow that is prediction estimate okay. 2$^{nd}$ thing is I will ask you that do an estimate of temperature of today using measurement upto today okay that is our xkk, see here what is x kk is the estimate of today is temperature using information upto today okay I can also post the session ultimately we are just estimating.

So I can say what is the best estimate of yesterday is temperature using information of today, what is the best estimate of temperature 3 backs using information upto today that is called if you can consider estimate you can and that is called as smooth estimated. So now you have collected more information see ultimately when you generate x k k it is estimate ultimately okay.

So if you have more information you can improve on the estimate okay, so smooth estimate are constructed, smoothing is done okay.

(Refer Slide Time: 51:06)

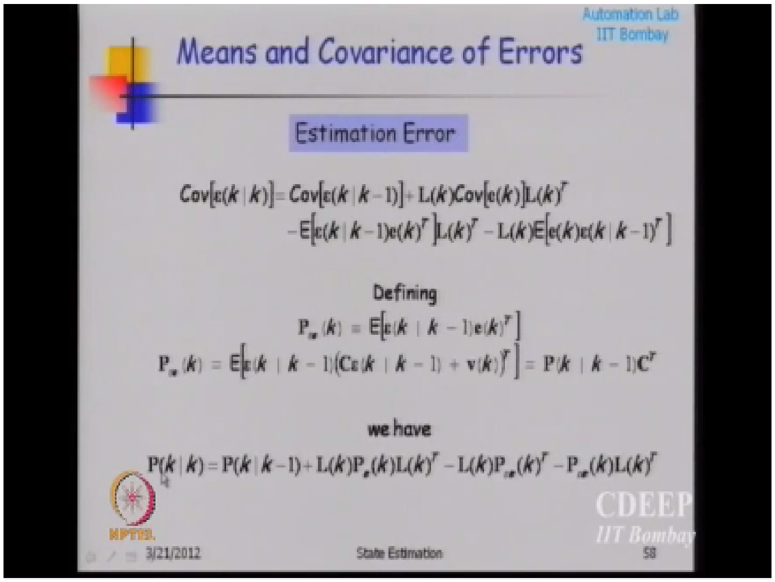



Just does this find out co variance of this ek, what is co variance of ek? You express it in terms of find out co variance of this ek in terms co variance of ε k / k- 1 what will it be? See Pk / k -1 c transpose + R, see what you will get will be like this, I am calling this ek is c this whole quantity into bracket transpose I am taking expectation of this, so you will get c.

Now this error and vk are uncorrelated so cross covariance between ε k / k- 1 vk is 0 and then that gives me this relationship okay. Each of the quantity stochastic quantity I have 3 stochastic quantities that I have worry about one is ε k k -1 other one is ε k k and the 3$^{rd}$ one is ek okay. So for all 3 of them I want to find mean and co variance okay. This is my preparation going on okay we will derive the observer after this.

So I am finding out now estimation error, first I will find it is mean value, now since I know mean of this use this relationship and find out mean of this. Why I am doing this after wards is because I have to go sequentially these are coupled equation and the I have to go one after other and the I find out.

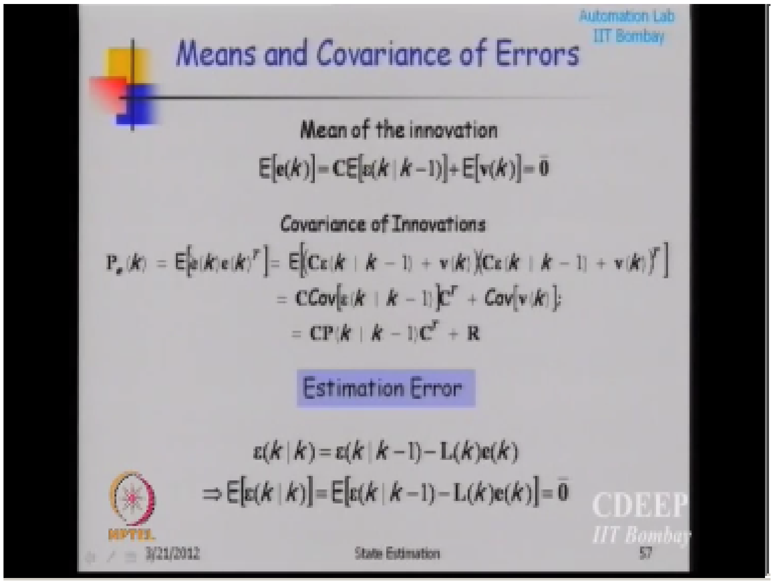
(Refer Slide Time: 53:16)




Kk okay is everyone with me on this, I am just finding out the covariance of this quantity and covariance of this quantity will turn out to be if you just do the algebra it will turn out to be. Now here there are some cross covariance terms that are coming up which I cannot neglect okay. So if I do the algebra you have to it can work out this little bit of algebra, if you do this little bit of algebra you will get this equation which says that, new covariance.

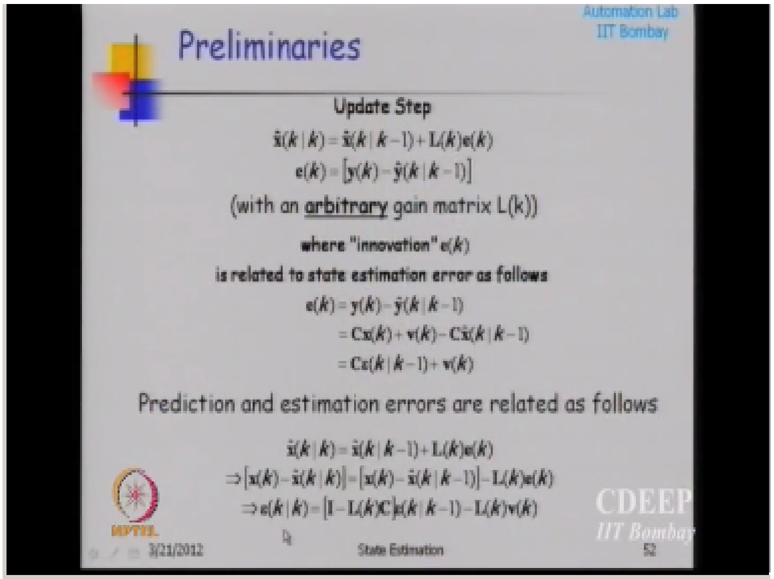
That the updated covariance is predicted covariance + Lk Pe is covariance of e itself and this is cross covariance between e and ε okay so what it tells you that the new covariance updated covariance is going to be function of Lk. How do you choose Lk will decide Pk / k okay? Now forget about the algebra what the covariance tells you spread. What kind of estimate that I want spread should be minimum, maximum, small, and large as small as possible?

I want to choose Lk, I want to choose Lk such that this covariance predicted covariance is as small as possible okay coming up to this point it just algebra you sit down with this equations patiently do the algebra you will get this equation I have given a all intermediate steps okay it is nothing they look little complex but you know just doing patiently doing the algebra and you will get these expressions okay yeah.
(Refer Slide Time: 55:46)



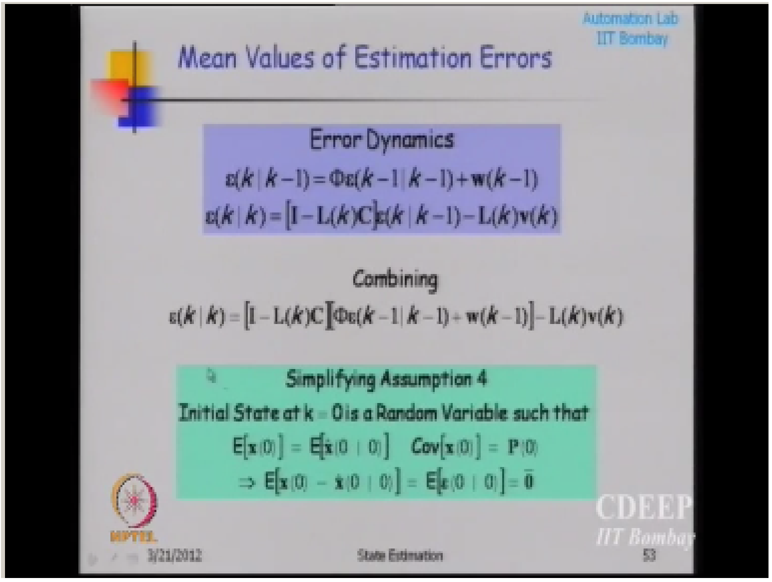

See there are this equation is there right k, k-1.lkek so you need to know both about ek and lk see why did I derive here pe here because I am going to get a term when I do χk, χk transform I will get a term which is ek,ek transforms I should be ready for that, that is why I have done this preparation which one.

(Refer Slide Time: 56:23)




Slide 50 they can be a type over which one we can derive it see here for this one minute no one but I did I subtitle for yk there.

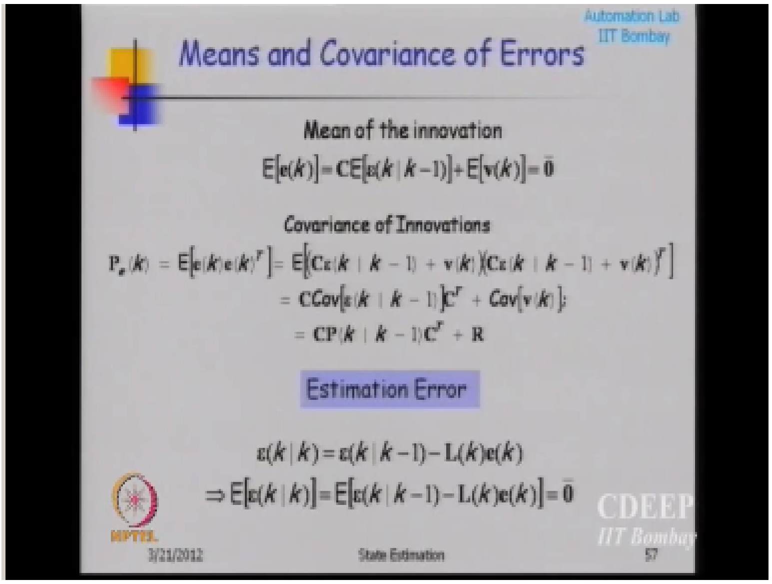
(Refer Slide Time: 57:20)




This equation right, have I made a higher here let me check oh yeah okay the way I derived this equation is simply by subtracting so you start with this equation yeah so I have just written this equation that is χk-1 oh here it should be k-1k-1 right okay this equation you are talking about this equation no these two are same equations see I can choose to use this equation or this equation.

So see here what I have done is I have expanded ek and just written bk here okay these two are one and the same equations here ek is written in terms of c χk given k-1+vk then you get this equation okay but you can choose  to work directly with this equation know so there is lot of algebraic tricks involved in this okay so this equation and this equation are identical actually they are not different equations so I could I could have processed using this equation I have decided to proceed this equation okay so they are not different yeah thank for pointing out.

(Refer Slide Time: 01:01:01)

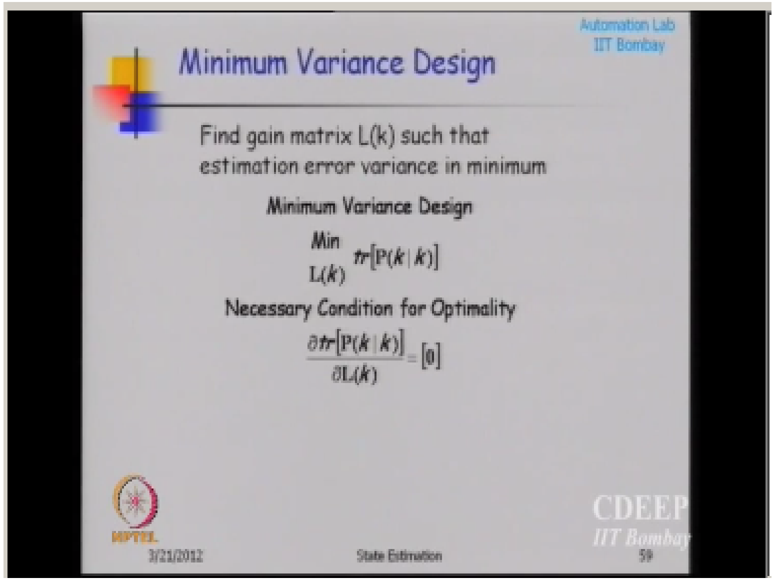



Yeah fine there are two different expressions for the same thing they are inter convertible okay is this file up to here last equation okay now what I want to do I want to device a minimum variance controller I want not minimum variance controller I want to device a minimum variance estimator I want to find out that gain lk which gives me smallest possible variance in the estimate of x okay I want to find out that gain lk which gives me smallest possible now what is the relationship.

Between the predict corrected updated covariance and gain lk that is given here okay I want to choose lk in such a way that this is as small as possible now this is a matrix what is the smallest what is the matrix as small as possible? Is this a special matrix is this a matrix which is positive to connect is the covariance always the defined matrix why is this matrix are positive defined matrix? Just look at the terms is this positive defined matrix covariance matrix just goes back

And look at the derivation initial covariance, covariance matrix p0 is always positive depend if p0 is positive definite you can show that p1 is positive definite if p1 positive definite p10 is positive definite if p10 is positive definite you can show that p11 is positive definite and it falls that all these are sequences of positive definite matrixes okay covariance is always a positive defined matrix it could be positive semi definite but it can never be negative depends so it is a positive definite matrix now, it depends upon how we choose this lk so question is how do I choose lk and maintain the positive definiteness of pkk okay.

(Refer Slide Time: 01:03:42)

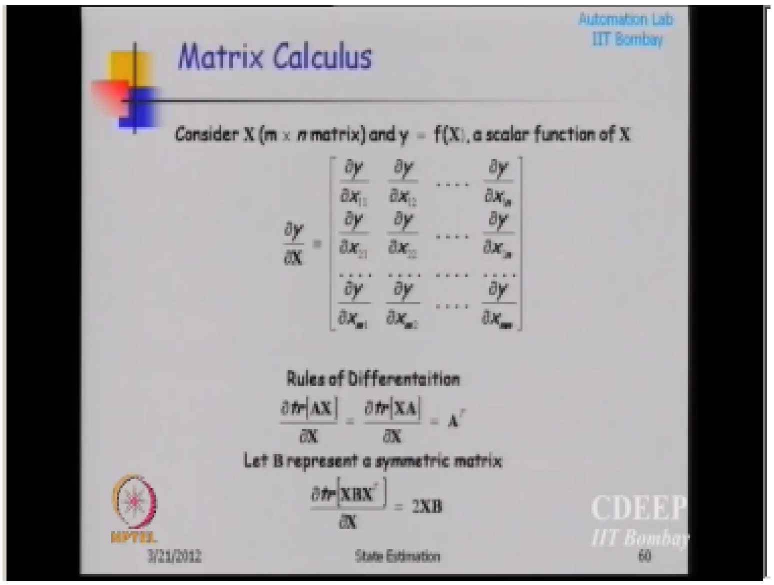



So now what I am going to do is I want to find out  estimation find out the gain matrix lk since that estimation error has minimum variance so I need some you know scalar quantity which defines the volume of this matrix I want to talk about large matrix small matrix okay let us assume that this is a diagonal matrix if the diagonal entry are large co variance is large so if diagonal entries are small covariance is small okay if this is diagonal matrix and to the positive definite matrix all the elements will be positive okay.

So trace what is trace some of the and also Eigen values some of the Eigen values what is the Eigen value of the positive infinite matrix are they always positive ? yeah positive okay so minimizing variance turns out to be equivalent to minimizing trace of this matrix okay how do you minimize the function with respect to some what is the necessary condition the necessary condition for optimality is derivative with respect to the objective function said equal to 0 so now what I need to do is to differentiate trace of pkk with respect to lk.
And set equal to 0 then I will get the solution okay then I will get the solution so once you I want you guys and girls to go back and do this algebra at home it is not just do not listen to thighs lecture just go back and try to derive these things okay first of all now, I need some intermediate result first of all you have to understand that trace of c+d same as place of c+b and trace of c is same as place of c transpose okay.

So I need this because now what I am going to do is I am going to take trace of pkk which is trace of this whole quantity on the right hand side okay so trace of this is some of trace of this plus that is one thing okay and some of the other quantity is see if you look at this and this matrix is transpose of this matrix but the traces are equal okay so I am going to use that so I need this relationship okay next thing I need is how to differentiate.

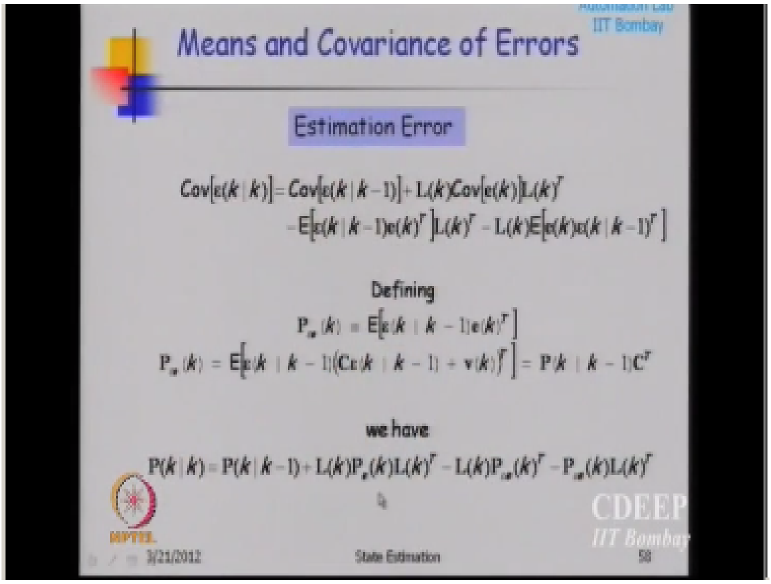
(Refer Slide Time: 01:06:46)




A scalar function of a matrix with respect to the matrix itself see what I am doing x is a matrix okay right now well I should not have used x is now not the state here I am writing the result mathematical result as a side thing x is some matrix and y is some scalar function of this matrix then what is ∂y/∂x that is defined like this okay I should have used some other notation x is not a great thing x not confuse on this particular slide x does not represent.

This state it is just a algebraic result which I have writing on the side okay now there are some rules of differentiation okay offers scalar function of a matrix with respect to the matrix itself okay so trace of a times x trace of a times x by ∂x, x is the matrix is a transpose you can prove this results with algebra you can prove this results okay use this definition take ax and then mathematically write down I tried it works out okay so there are also results for x transpose bx why do I need all these look here.

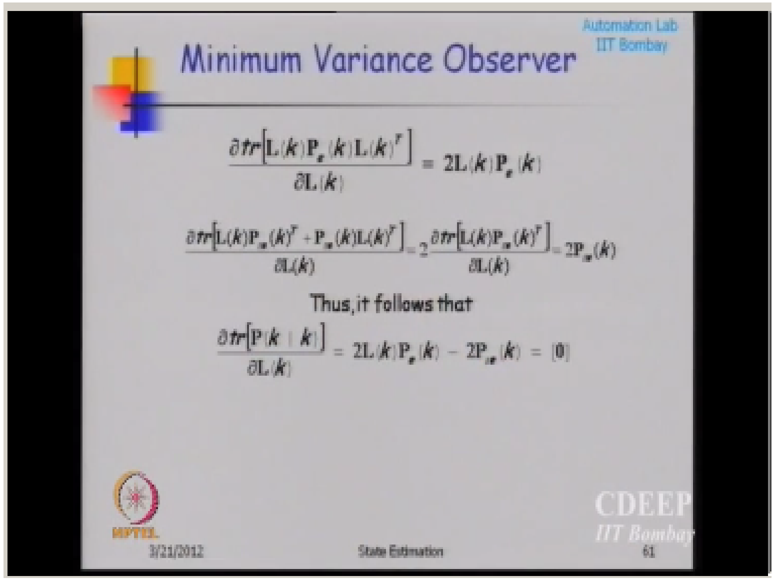
(Refer Slide Time: 01:08:27)




I have a term which is matrix l into this transpose so I need to differentiate this term I need to differentiate this term right I need to differentiate these two terms what is derivative of this with respect to lk yeah if he said it 0 pk given k -1 does not depend upon lk only these three dependent so I have to differentiate this term okay.

So differentiate these three terms I need these two results these two are algebraic results one is you know when you differentiate trace of ax or xa with respect to x you will get a transpose then you differentiate xb transpose with respect to x you will get 2xb okay this is just algebra the final result looks exactly same as scalar differentiation except here you will get a transpose okay otherwise it looks almost you know close to I am going to use this and differentiate okay I have to differentiate first this term

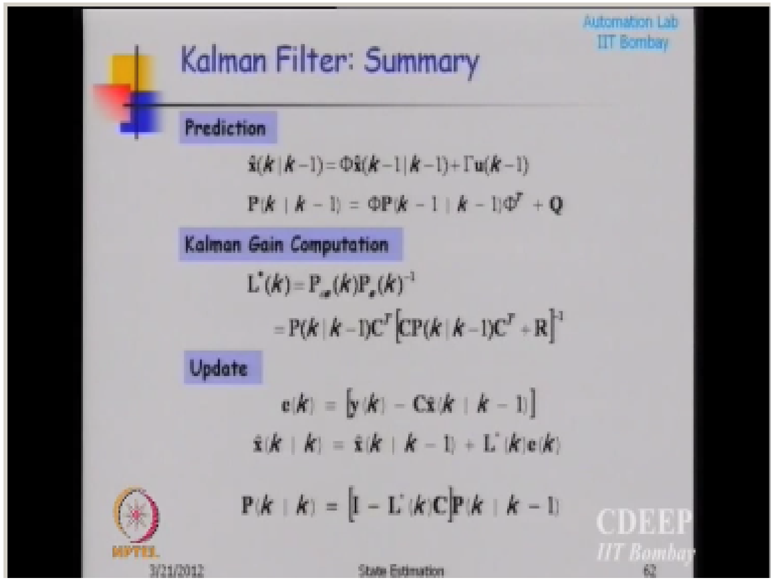
(Refer Slide Time: 01:09:47)




I will get 2lkpe okay I have to differentiate this term and from this I will get 2 here is everyone 2 results is it okay just I am using that algebra I showed you differentiating a scalar function of a matrix with respect to a matrix what is a matrix here l okay I am differentiating pkk with respect to lk so finally when I do this I will get this result okay.

So what is the optimal gain is this l* which is p χek.pee$^{-1}$ okay this is my optimal game if I use this value of gain I get smallest possible variance pk given k okay this was the fundamental this is similar result root by and then what is the optimal covariance turns out to be this smallest covariance this is the smallest covariance that you can get you cannot reduce covariance below this okay this is the optimal estimator okay

(Refer Slide Time: 01:11:35)




And so let me summaries kalman filter this is my prediction step and then I do prediction of the covariance see I am working with the stochastic process I have to keep updating mean covariance okay so prediction step is this so this is predicted mean this is predicted covariance okay kalman gain computation is you know this p $\chi$e.pe$^{-1}$ we have computed these terms earlier and kept them aside I am just substituting them here okay we have done this calculation just look at the slides we have done this calculation.

So this is how you compute the kalman gain so if you know pk given k-1 you can compute Kaiman gain and if you know ϕ if you know r and you know q so what I have achieved see I started by saying I want to derive the gain by systematically using stochastic information about w and v I have done that my gain calculation is based on q and r how it is depended it is depended on q through pk given k-1 is appearing here pk given k-1 is function of q okay lk is the function of r okay.

So I have systematically used q and r information to find out so this is the celebrate Kalman filter I do an update and then I update covariance okay this is updated mean updated covariance so you start with the initial guess x00 and this is the recurrence procedure by which you construct subsequent optimal estimates for linear timing variance is linked these are the estimates which give you minimum covariance which they is known better estimate and you can construct they are lot of properties I mean this particular development in 1964 or 65 about a time I was born you know guess gave rise to results.

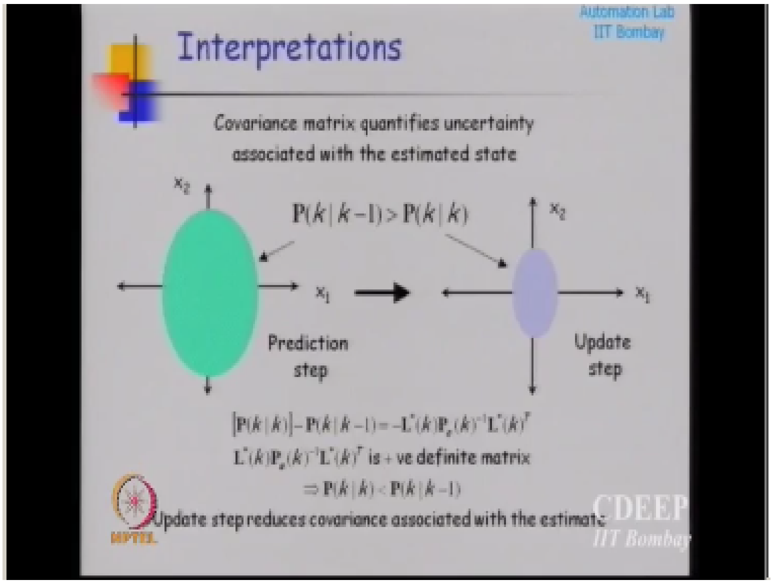
And it has resulted into huge number of things that you just do not know what arte the things that we use this thing is used in as I set image reconstruction it is used in you know under water drilling it is used in see you have a problem in oil way drilling you have some measurements coming from the you know the drill as some measuring instruments okay it will measure some humanity and some other things I do not know what exactly it measures then I have a model for the reservoir okay.

And some those measurements you construct the state what is the state what is the reservoir down there okay what is you know how soft metal is on this side and how soft the metal is on the next side you can estimate you have a model you are data coming you fuse data in the model that is what you doing here see this is my model based estimate my data why is coming here this is my predicted estimate I am fusing data with the model in this okay a wonderful algorithm why this is wonderful?

Because this is recursive so, when gauss initially worked out on this square estimation he worked out on a batch of data okay batch of data is not it is brilliant solution but you know in a computer control system where data is continuously coming data sixe keeps increasing you cannot work with batch of data using a recursive solution so this is a recursive solution go and estimate plus a correction gives me an new estimate.

So that was something below the land mark because this could be used in the computer okay it just says that updated means is five times the previous mean plus this and corrected mean is updated predicted sorry this is not updated this is predicted mean and corrected mean is plus correction, correction is okay very, very powerful algorithm which.

(Refer Slide Time: 01:16:01)




So now look at its interpretation and all kinds of things in the west lecture what I will show you is that every time there is a covariance reduction so every time you getting a better and better estimate and so on okay.

**NPTEL**
**NATIONAL PROGRAMME ON**
**TECHNOLOGY ENHANCED LERNING**

**NPTEL**
**Principal Investigator**
**IIT Bombay**
Prof. R. K. Shevgaonkar
Prof. A. N. Chandorkar

**Producer**
Arun Kalwankar

**Project Manager**
M. Sangeeta Shrivastava

**Sr. Cameraman**
Tarun Negi

**Sr. Online Cameraman**
Sandeep Jadhav

**Digital Video Editor**
Tushar Deshpande

**Technical Assistants**
Vijay Kedare
Ravi Paswan