Optimal Control Guidance and Estimation

Prof. Radhakant Padhi
Department of Aerospace Engineering
Indian Institute of Science, Bangalore

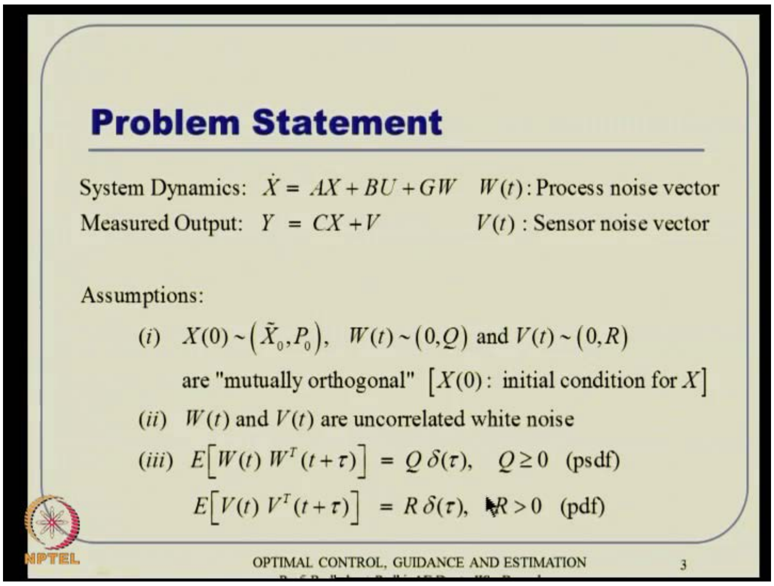
Module No. # 12
Lecture No. # 29
Kalam Filter Design- ll

Hello everybody, we have started discussing on kalman filter for last couple of lectures. We actually have given some overview and then followed up with some basic concepts of random variable followed by kalman filter derivation in continues time linear time invariant system that platform actually.

(Refer Slide Time: 00:21)

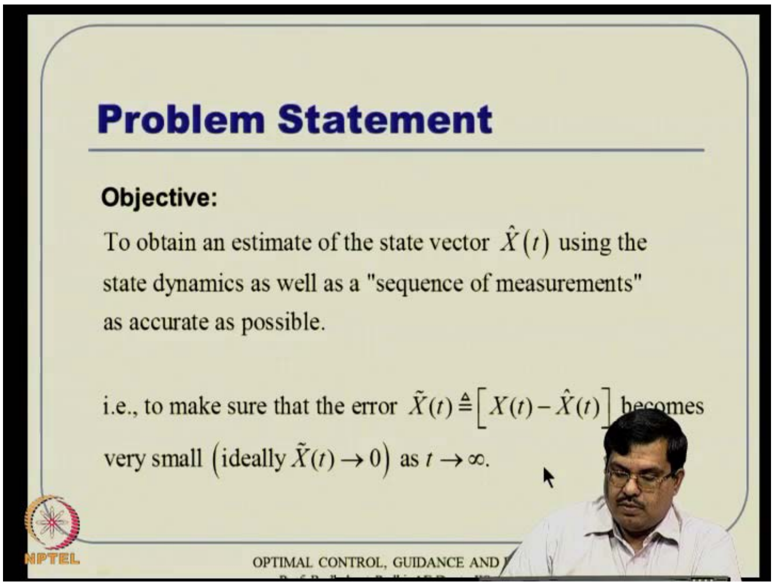
So, we will develop further and our motivation is to go towards external kalman filter, and if possible uncentered kalman filter as well, that is what people use these days variably, that is what to be double up toward that. This particular lecture will again have a very quick overview of what we discussed in the last class followed by the discrete time domain a derivation actually. And then we will continue further on those lines, ultimately the idea will be to merge discrete time and continues time together. Alright, so let us get started, a very quick overview of kalman filter design for linear time invariant system and as against continues time domain of whatever we have discussed in the last lecture.

(Refer Slide Time: 01:25)



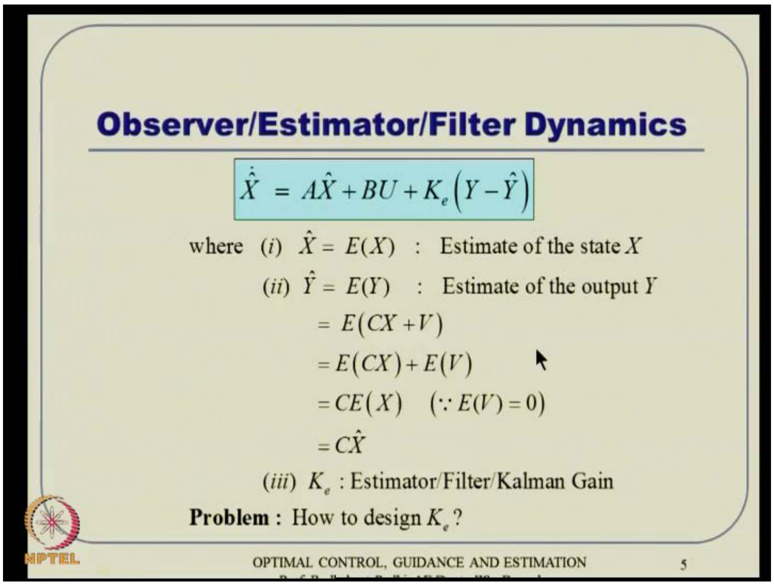
And the system dynamics was something - I mean - consider here is something like this Xdot is AX plus B U plus G W Y is C X plus V, W and V are continues time process and sensor noise is respectively, and also at some assumptions that initial condition are described something like a mean value and associated co variance matrix, and followed by all these assumptions that W and V are uncorrelated white noise. And then they are also mutually orthogonal all these X naught W and V are mutually orthogonal think like that. And using this relationship somewhere down the line we are able to derive this continues time kalman filter, so how do you do that, something like this the objective was to estimate this X hat of T (()) using the system dynamics as well as the sequence of measurements as accurately as possible that was our aim.

(Refer Slide Time: 02:12)



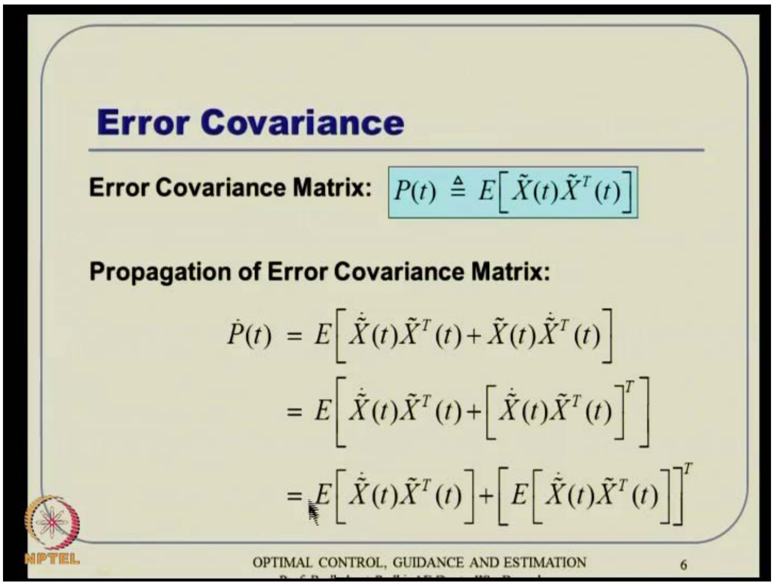
That means, the error of estimation X minus X hat of T should become very small or ideally it should go to 0 as T tends to infinity, so that was our objective.

(Refer Slide Time: 02:37)



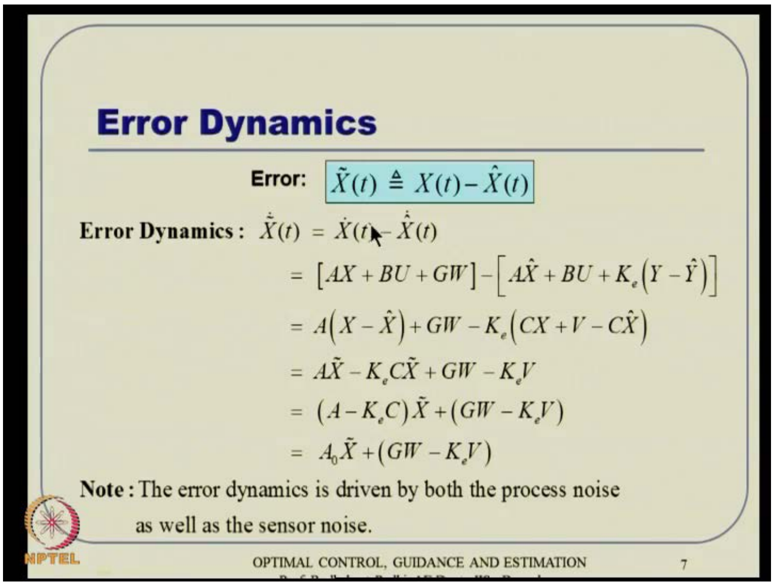
So, how did you do that? We had a estimate of dynamics which you took it in this form, X hat dot is A X hat plus B U plus K E times Y minus Y hat and where x hat was defined as expected value of X and then Y hat - I mean - it turns out that is nothing but C X hat because E V is 0 - I mean - expected value of V is 0, so the whole point is, now how do design this kalman gain K e, but then we had some derivation like this.

(Refer Slide Time: 03:07)



First we will define something like a process - I mean - sorry this P of T which is nothing but the error covariance matrix, so this was you find like that and we wanted to have study a dynamics of that how it all, where it goes and think like that actually. So, we had this P dot derivation which is nothing but expected value of all that, because expected value and derivatives are both linear operators they compute, so the derivative goes inside and it follows like this, and turns out that the this P dot can be expressed as something like this, plus a this entire something while transpose, so then with (()) what is this - this quantity?

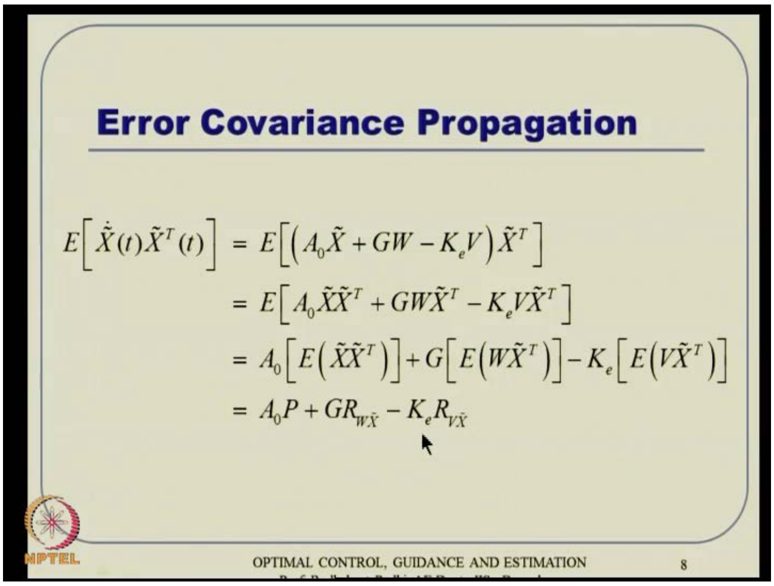
(Refer Slide Time: 03:47)



So, that quantity is a - I mean - for knowing that we know this there is a term called X tilde dot. So, for knowing X tilde dot we have to go back to X tilde definition, and then form this definition, this X tilde dot comes out naturally, we have X tilde dot is X dot minus X hat dot, which put back the dynamics, and X hat dot is the observer dynamics put it back and then try to simply further, and turns out the this is a something that what you define a minus K e c is a dot X tilde plus this additional quantities G W minus K e v.

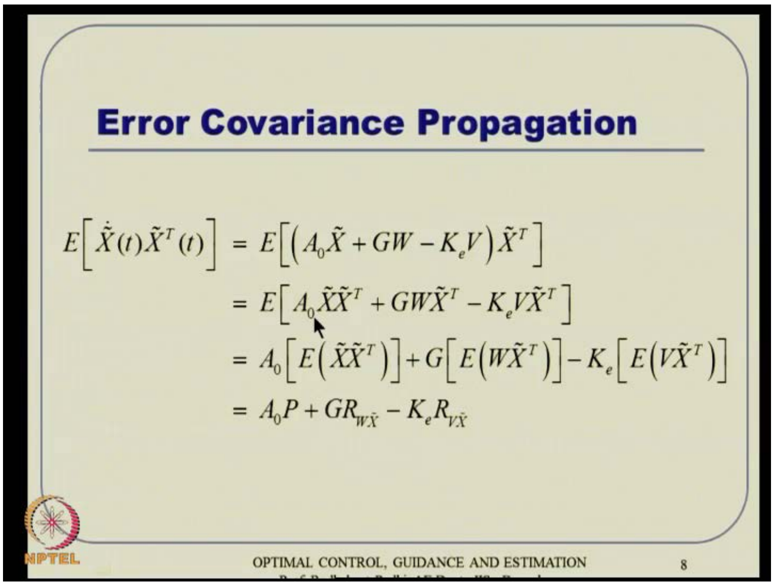
So, obviously the error dynamics gets affected by both process noise as well as sensor noise basically. Then, this term is ready now but our aim was not that, our aim is to get some value something like this.

(Refer Slide Time: 04:39)



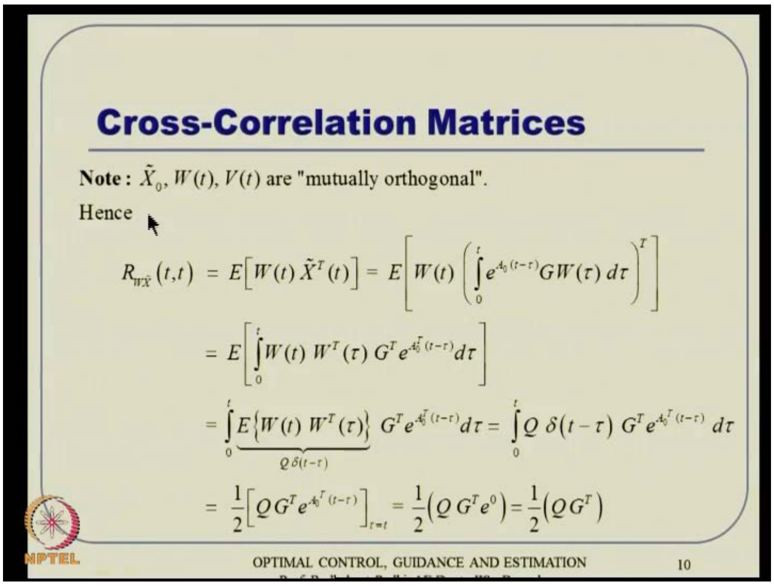
So, then it all, so what is this one, this one we expected value of that, and then we try to expand the algebra inside and then invoke this idea of that expected value is a linear operator, so we can separate it out. And then it turns out that this quantity by definition is nothing but P, however, this quantities we need to still evaluate - expected value of this guides still to evaluate actually. So, where you heading to - I mean - this quantities for knowing these we need to have a solution of X tilde dot actually what we have got, so far is X tide dot. So, we need to get a solution for that, but then you can think that this X tilde dot equation what you have here is something like homogenous part and time varying input.

(Refer Slide Time: 05:20)



So, if this is like that we know the solution and the solution is coming from linear time invariant system theory directly, then this solution can be represented something like an exponential term with a initial condition plus this integral term of this - convolution integral - sort of thing where 0 to t e to the power a naught t minus tou all these actually this is nothing, but the entire thing is nothing but a time varying input. So, then again integration is a linear operator, so we can separate it out and then we can evaluate because - I mean - this is the solution of X tilde t, but this is what is our aim? Our aim is to evaluate these quantities.

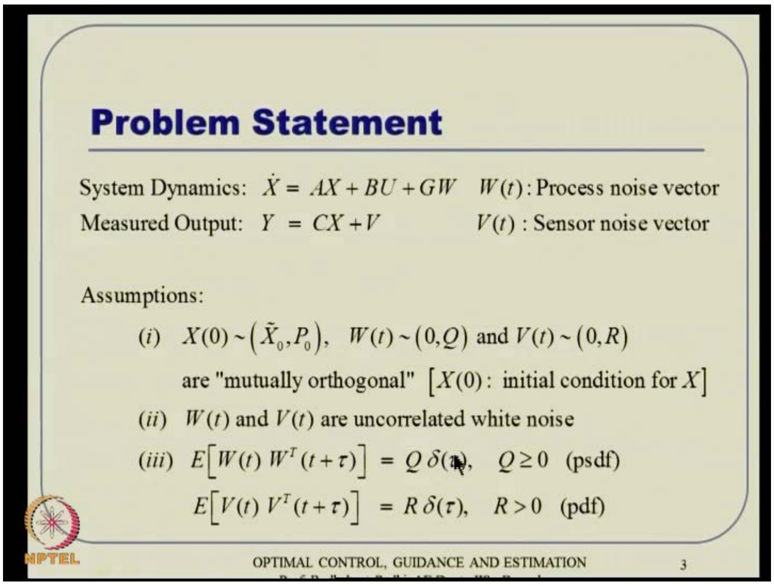
(Refer Slide Time: 06:09)



So, we go back and substitute the term for that quantity R of W X tilde, so by definition is nothing but expected of W and X tilde transpose. Now, we have got a expression for X tilde I think, so we put it back there, and in turns out that because of mutually orthogonality relationships and all everything else will go the v and X tilde dot will go with as long as it is multiplied V W actually.

So, only term that will remain is W term itself, and again you can accept this linear operator concept, and it was this Winside. And then here it turns out that W times W transpose appears, and this again, this expected value will go inside the integral, and that expected value of this thing by definition that expected value of W W transpose is nothing but Q.

(Refer Slide Time: 06:56)

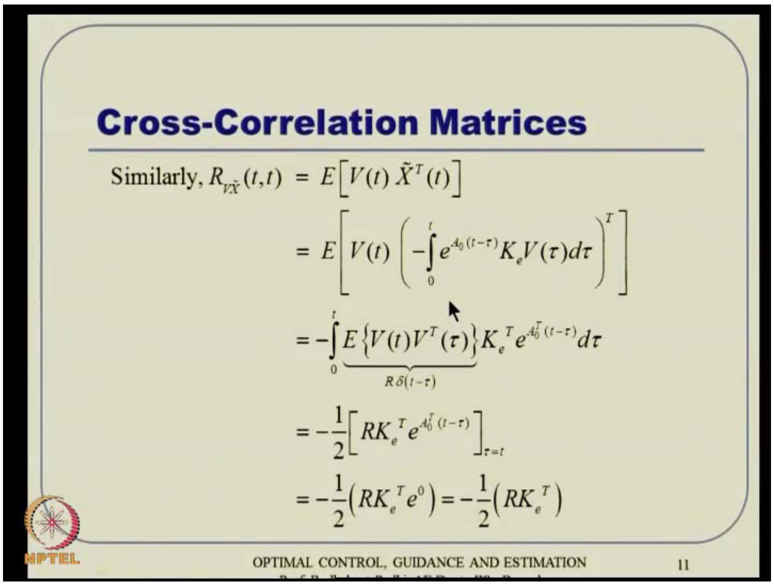


So, because of that it turns out that, this is nothing but delta operator will coming here, and we also - I mean - there was a small (()) their last time, but anyway this turns out that if you have a delta operator sitting here something like integral a to b f of t and multiplied by a delta operator with is, will let say, t minus tou dt sort of thing or t minus let say C (()) alright. Let say t minus c into d t, this turns out that if c lies completely inside the interval of a and b, then it is nothing but f of c, and if it is somewhere either a or b, then it is something like half of f of a or half of f of b.

So, this one is provided t e is completely inside a b, this is if, sorry, if its c - one second - if its c is completely inside a b, and if this one if c is equal to a what happen that, so this is actually c is equal to a, and this one if c is equal to b; that means, as long as c is completely inside the interval of this integral, so that takes the form of just as y - I mean - f of c, but if it is somewhere on the boundary then half term will come into picture.

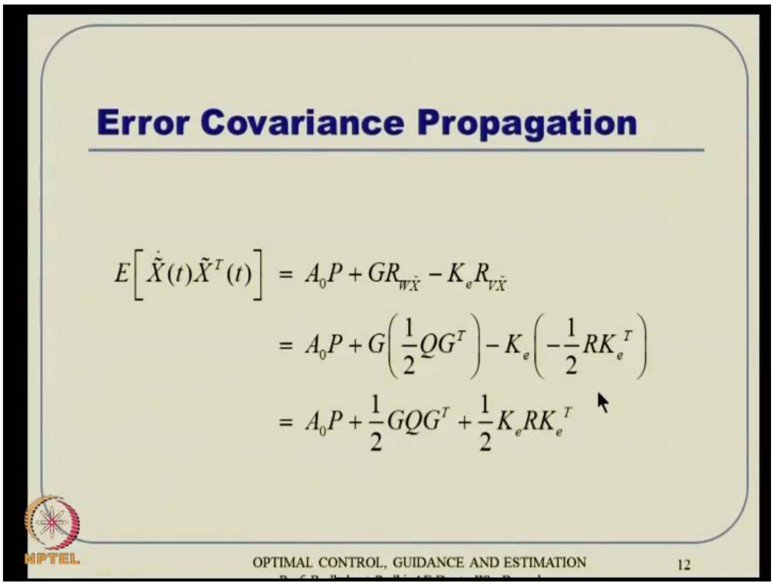
Because of that this interval, what this integral what you see this half term here, but tou is getting evaluated at t, because t happens to be one of the - I mean - limits of the integral. So, then look at this is nothing but get - I mean – now, it is e to the power 0 goes to equal to tou, then E to the power 0 is nothing but identity and hence you get something like this. So, you get, now it a kind of deterministic quantity for this term cross covariance matrix sort of thing.

(Refer Slide Time: 09:39)



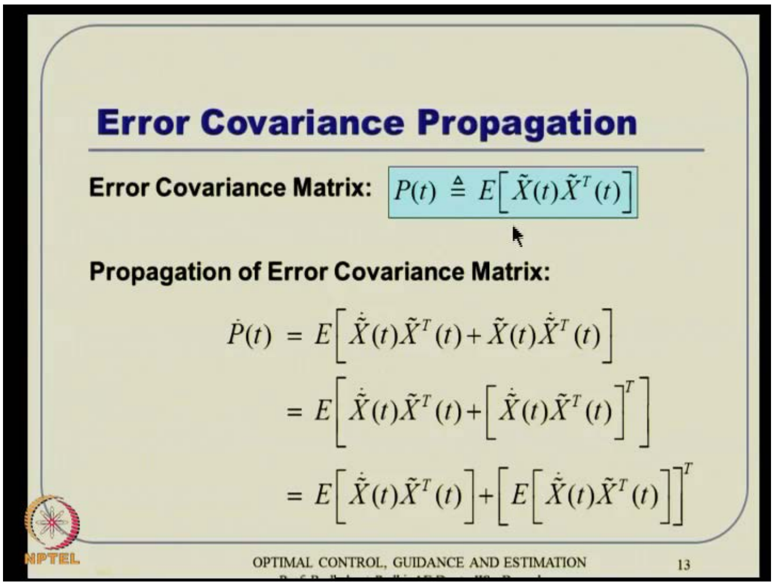
And then forward, this quantity we can get follow the similar line, what instead of this - I mean - only returning this now, it will be only this quantity with a negative sign of course. So, this algebra follows exactly similar to that again, we will have a delta function there and it turns out that something like minus r into K e transpose.

(Refer SlideTime: 10:00)



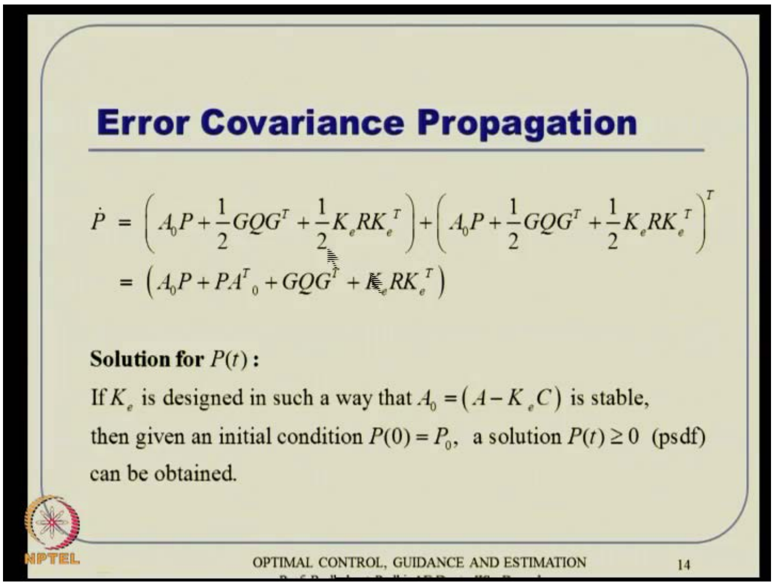
So, then note down, then now we ready to put everything into place, so our motivation wants to evaluate this first. So, this term was already there, now about this two terms we got something like r w x tilde is something like this and r v x tilde is something like this.

(Refer Slide Time: 10:19)



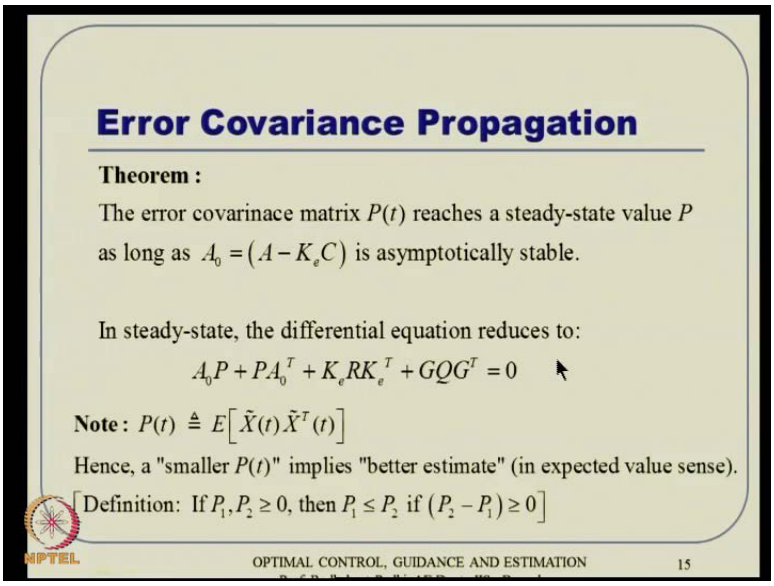
So, now this expression turns out to be something like that, so going back tell our P dot was something like this term plus the same term transpose.

(Refer Slide Time: 10:24)



So, hence P dot is nothing but same term plus same term transpose - I mean - this term plus for this exact same term transpose, if you do the algebra and then combine this terms, it turns out to be something like that. Now, here is a problem, where we still need to design K e what we need a solution for p also for that.

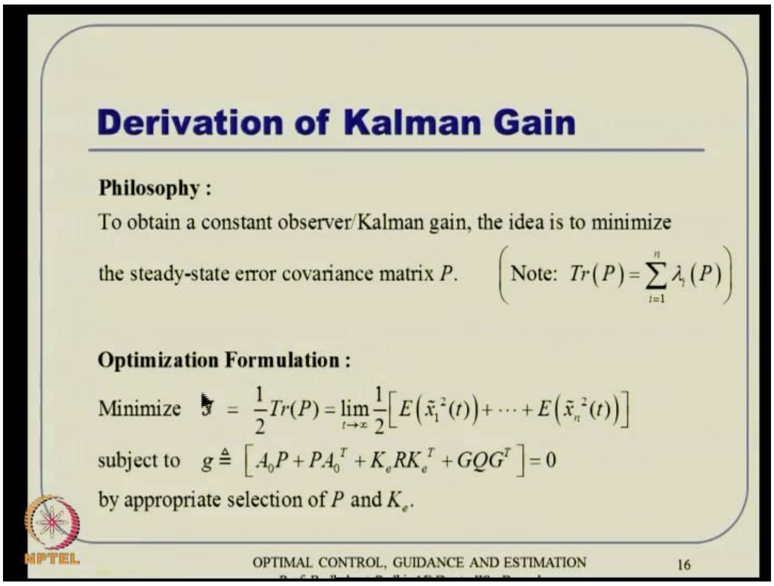
(Refer Slide Time: 10:44)



So, here we have to go back and X (( )) theorem sort of thing, it tells and steady state this P dot is going to go to 0 provided this quantity a naught which is a minus K e c stable actually so and what you are interested in is a positive some definite solution for P of t, so this turns out to be like this.

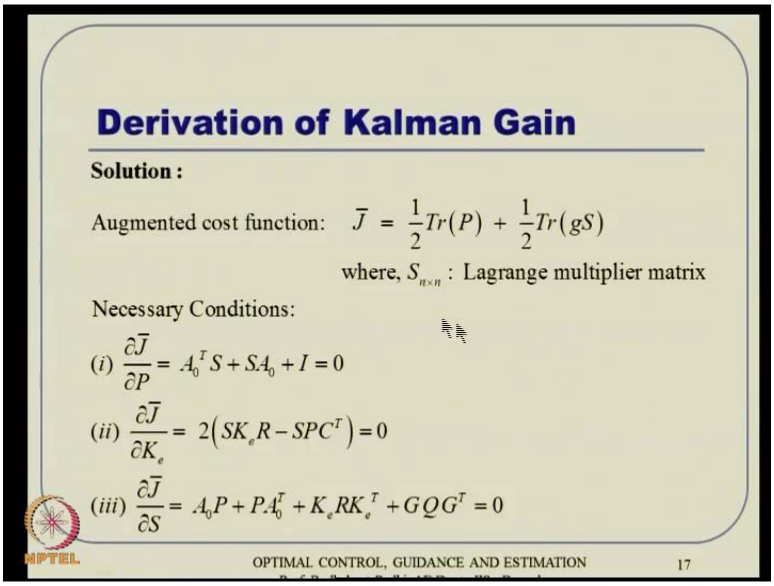
So, essentially, what it mean is because P of t something by d - I mean - P of t is something like this, but it also mean it is smaller than P better actually - better the estimate. So, essentially what we are telling here is, we want to minimize p, so objective this constant equation.

(Refer Slide Time: 11:27)



So, we have to formulate some sort of optimization problems where we have to minimize j is half of that, so objective of this constraint for this is a matrix equation constraint, this matrix norm trace is also kind of norm.

(Refer Slide Time: 11:44)

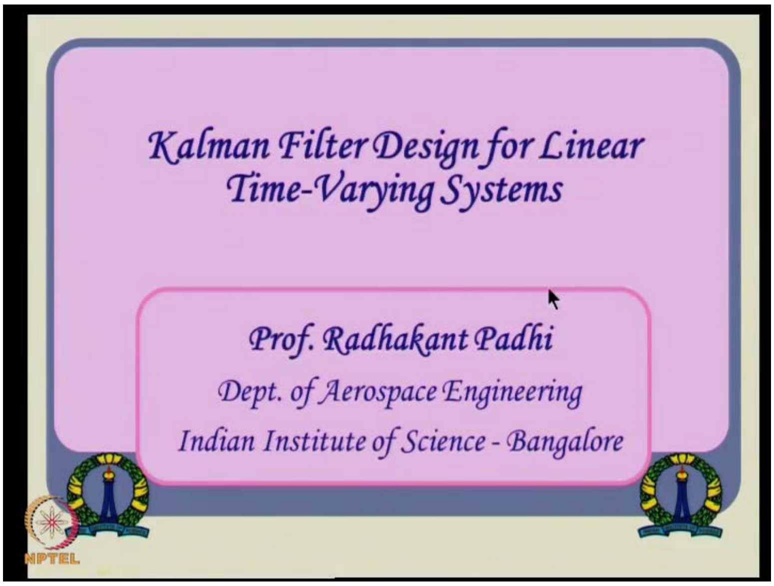


Alright, so then what I told is, the formulation has the solution turns out to be something like that you formulate an augmented cost function and then take all this partial derivates with respect to P K e as well as this lagrange multiplier matrix s basically. So, first all this three constraint equations is to be satisfied together, and hence, if you solve it one by one, it turns out that K e is nothing but P C transpose R inverse actually, so that we got an expression for K e.

Now, how about is P is, P can be obtained as a solution of this, because all what you are doing here is getting this three equation satisfied together basically. So, it turns out that this equation must be satisfied, after putting all these expression for k e and hence it - I mean - this nothing but what is called as filter Ricatti equation actually. So, sincerely we need to solve this equation and then evaluate this kalman gain like this, and then put it back in the observer equation that we know actually, that was the whole development in the frame work of continuous time lti.

But it turns out that certain - I mean - in variably we have assumed one small thing here which has the big implication, that the measurement equations we also assumed, there continuous time and unfortunately the continuous time measurements are typically not feasible, so next people thought how got going to discrete time actually.

(Refer Slide Time: 13:11)

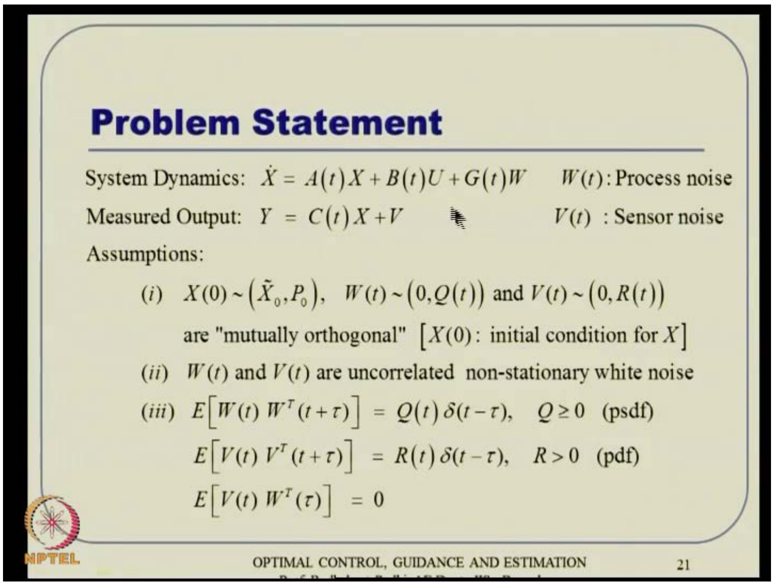


Because if you have a continuous times system dynamic equation, we can always discretize it in using some discretised numerical equation - I mean - numerical integration formula and think like that. So, that will give us a platform for discrete times system dynamic whereas the discrete time as sensor measurement are still available with us actually directly, so that makes if the platform some sort of a comfortable actually.

So, then that is our idea that we wanted to see kalman filter design for linear time - I mean - discrete time thing, but before that there is one more assumption that - I mean - we had it here, in the entire beginning - if I go back - we also told that this is nothing but the linear time invariant system; that means, a b and - I mean - C all that what you see these are all not varying with time.

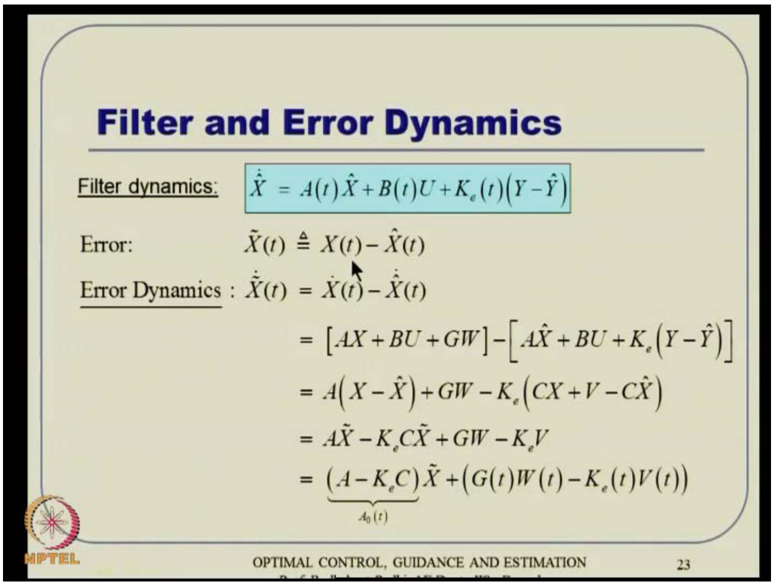
Now, what if they are time varying things, so that is the natural generalization before you go to discrete time domain. Alright, so then let us see that, this development is very parallel to what we had done for linear - I mean - continuous time l t I system, so let us see what is that derivation for that.

(Refer Slide Time: 14:29)



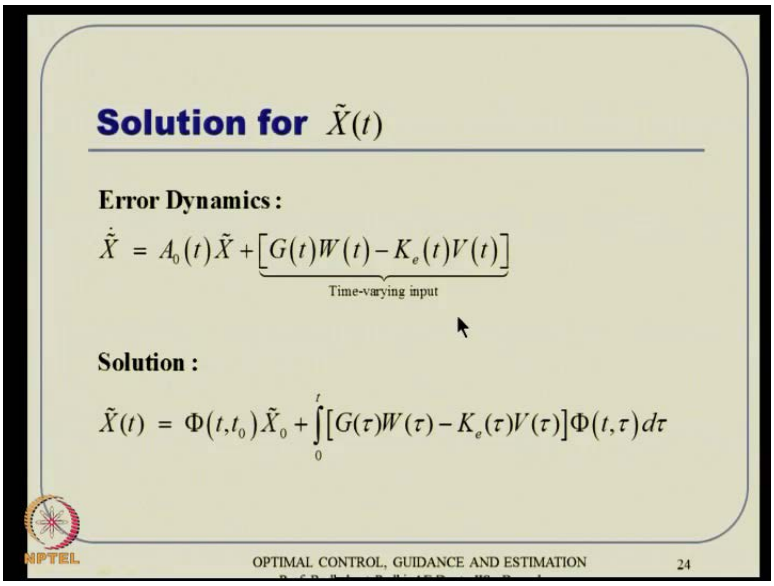
So, the problem definition is almost same what we seen before, the only difference is all the system are dynamic matrices A B G C and all that they are all now time varying matrices, they take different values different point of time, so then what now actually. Alright, so the objective remains exactly same again, we want to have some sort of an estimate for which the error should go to 0, X T goes to infinity.

(Refer Slide Time: 14:57)



So, we again had the same filter dynamics, and then we have this error definition x tilde something like this, then we have this error dynamics we will exactly derive similar way, the only thing that we need to keep in mind is that the all the matrices that we are talking now is time varying.

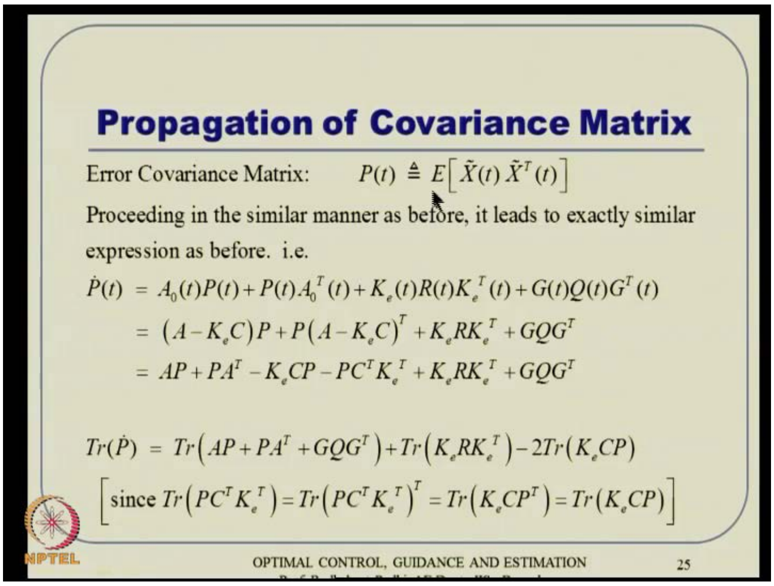
(Refer Slide Time: 15:21)



So, then it will land up with the same error dynamics, but the difference here is this term is actually time varying, so we cannot excite the solution in the form of exponential actually. However we know that even if there is a time varying system dynamics, but still it is a linear system, then we can have solutions of this form where phi comes into picture and this phi turns out to be something like state transition matrix. So, using this state transition matrix concept we can still write the solution something like this. And then we go back to the error covariance matrix definition P of t is nothing but that. So, exactly is proceeding very similar manner and that is one of the reasons why I wanted to review that last lecture material.

So, we again go back to the definition of what is P and then P dot is nothing but expected value of derivative all that the derivative, this derivative times that plus this time derivative all sort of thing, we can derive everything and then turns out to be something like this.

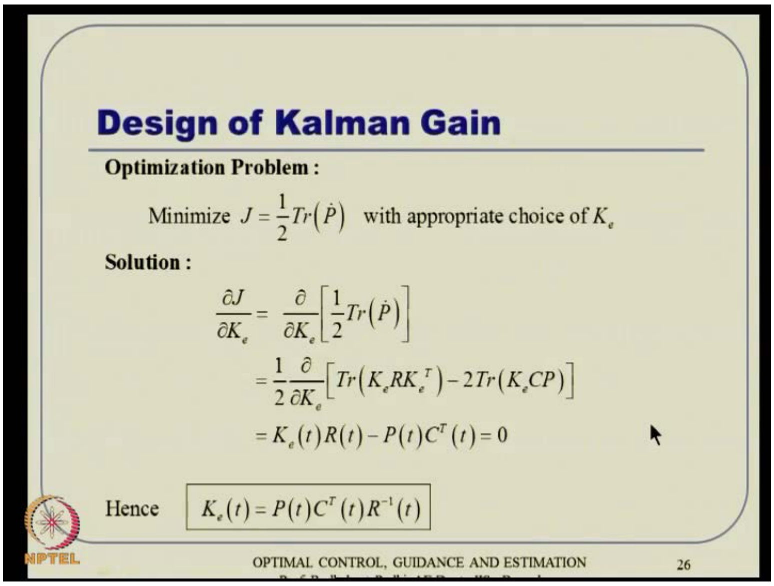
(Refer Slide Time: 16:01)



Okay, now what you can say is, we cannot excite the thing that on steady state things will go to 0, that is not possible, because it is all happening in a time varying sense the depends again the way these matrices vary and thing like that. So, we cannot excite this condition of asymptotic stability and think like that, even T goes to infinity we cannot claim that V of T should go to 0 - that may not happen here. So, if that does not happen then the next best thing are probably comes to mind is, how can we do something, so that the rate of change will be minimum actually; that means, P dot will turn out to be minimum.

So, then we want to minimize P dot and hence we see what is this trace of P dot actually. So, trace it turns out to be something like this, and since trace of this quantity is also trace of the same quantity transpose or same quantity transpose turns out to be something like this, the transpose will take a reverse sequence. So, it will be K e transpose whole transpose again K e c transpose whole transpose is C and then P transpose coming here, and remember, P is a symmetric matrix, so P transpose is P, so that turns out to be like this, so this affect has been utilized somewhere here actually.

(Refer Slide Time: 17:53)

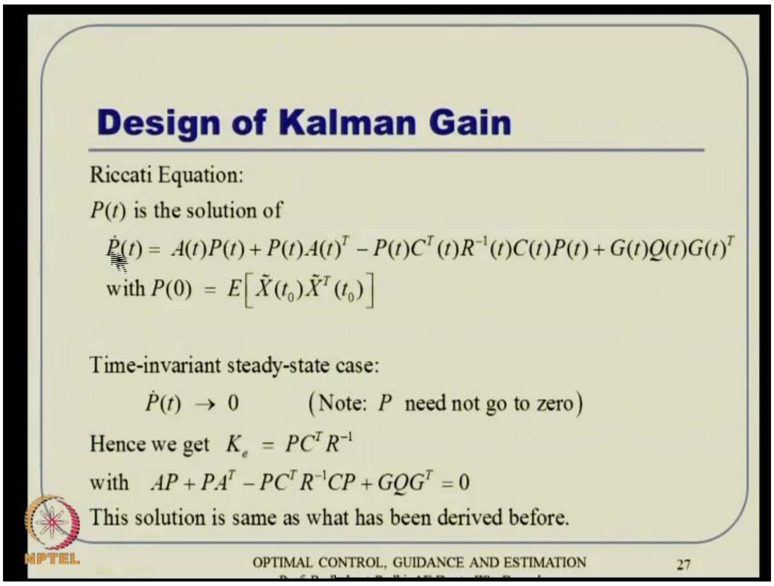


Now, the problem turns out to be like we do not want to minimize P for say, but we want to minimize P dot, so minimize this quantity j is a half of trace of P dot actually, and minimize with respect to what? We minimize with respect to K e basically. So, then the necessary condition is del j by del k e has to be equal to 0, and del j by del K e if you do this operator then it turns out that P dot is this expression, because these two quantities - I mean - P dot, trace of P dot is this quantity, and trace of P dot happens to be - I mean - this P dot happens to be this minus that, if you substitute back here this expression, sorry, this trace of P dot if you simplify the algebra and think like that it will turn out to be something like this quantity basically.

So, this quantity if I have this minimum - I mean - derivative with respect to K e then I can take it inside, tell this expression is nothing but this expression now, and this expression if it has to be equal to 0; that means, this if I solve for K e, K e takes the form of - I mean - this remember K e times R is equal to P times c transpose.

So, R is in the right hand side, so you have to take right hand side inverse only we do not have a choice of left side. So, K e turns out to be something like P C transpose R inverse again exactly the same expression what you have seen before but all these are time varying now basically. Now, that P dot what we are talking, but this is a P term where P comes from the solution of this P dot equation, with initial condition that we know basically t naught is nothing but P of 0 is nothing, but expected value of this quantity, where this is actually available with some just formula and all that.

(Refer Slide Time: 19:46)

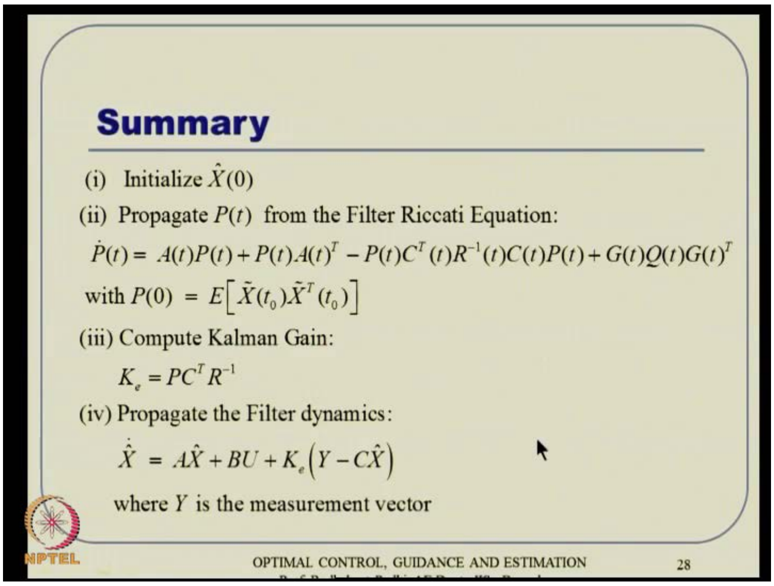


But anyway, so that is - I mean - P not is available, so P dot is available those of that utilizing that we should be able to get a solution for P of t actually. So, that is the approach, the expression remains same but p cannot talk about a steady state Riccati equation solution, but it should be a solution of this differential equation with this initial condition.

Now, what happens to the time invariant steady state case that is a special case sort of thing, then if it is time invariant P dot has to go 0, and if P dot has to go to 0 this expression remains same, but this expression turns out to be nothing but the same algebraic Riccati equation for filter design. So, essentially what it means the solution is same as what we have derived before.

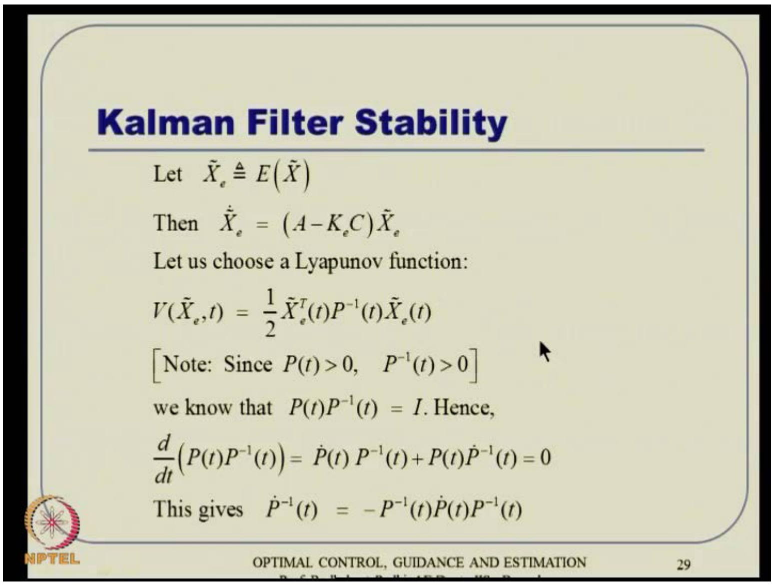
So, the only difference is the riccati equation is to be solved from this initial condition and this differential equation that is about time line. So, summary sense we have to have - I mean - you should initialize this state X at 0, and then we should propagate P of t from filter Riccati equation with this initial condition, compute the kalman gain and then propagate this filter dynamics, so that is the reason like that.

(Refer Slide Time: 20:54)



Alright, so everything good, but there are other practical difficulties and think like that which I encourage you to read some text books and all that, for mainly you coming from some numerical things as well as the difficulty for guessing some good initial condition and thing like that. But as long as it linear system dynamics and all it should not be too much of a concern, eventually everything will converse actually that is not a problem at all.

(Refer Slide Time: 21:54)

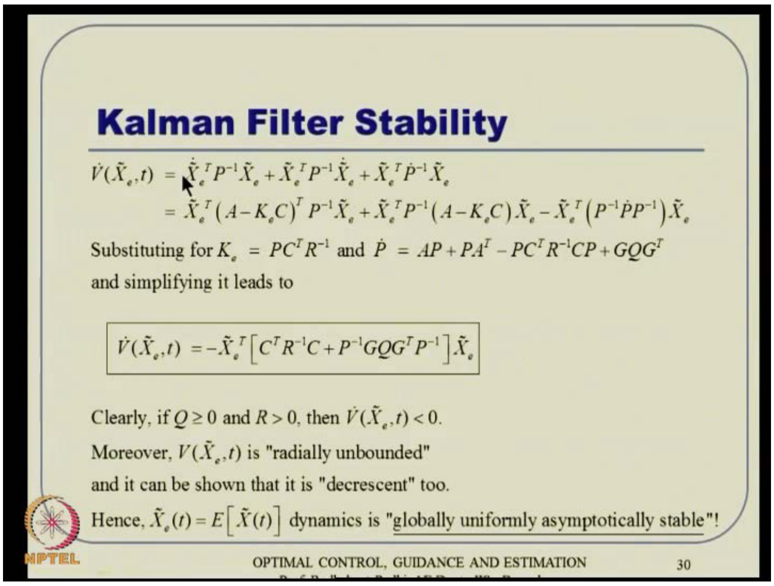


Alright, we will continue further and then tell what about this filter stabilities basically, because everything is changing in time I do not know what will happen and think like that. So, then we will define this X e tilde for algebra simplicity we introduce this new notation x e tilde is expected value of X tilde, and then X e tilde is - I mean - this expected value of X tilde dot and we have derived this before that this d by d t of this quantity turns out to be something like this actually.

So, this is available with us this expression we know basically, and then we tell will we will select a Lyapunov function which is nothing but this kind of this form and again remember Lyapunov function can be chosen anything as long as the satisfies certain properties like a positive definite and all that. So, this is the positive definite function, because p is positive definite P inverse is also positive definite, so we can select it that way. So, what we also know that in the derivative, so the Lyapunov theory will tell you analyze - I mean - select a positive definite v and then with respect to that selection v dot has to be a negative definite something like that.

So, when you take v dot, remember p inverse is also a time varying quantity; that means we will also need a derivative for P inverse actually - d by d t of P inverse. So, that is why till if the exciting this quantity P times P inverse is identity, we take derivative of both sides, and this derivative because identity is constant matrix this turns out to be 0. So, when you solve for this, then if you solve for this quantity, then it turns out that this is the expression actually for P dot inverse.

(Refer Slide Time: 23:43)



So, we go back to V and then tell what is V dot now, V dot turns out to be all these three expression first term, this rest of the terms - I mean - X e dot transpose and then this quantity and this these two quantities times X e dot and then really quantity dot actually P inverse dot. Now, P inverse dot is available, so this is our p inverse dot, and this X tilde dot is available that is what it is, and X tilde dot transpose it transpose of the entire thing which is reverse transpose sort of thing, so let see X e tilde transpose coming first and other thing coming later actually.

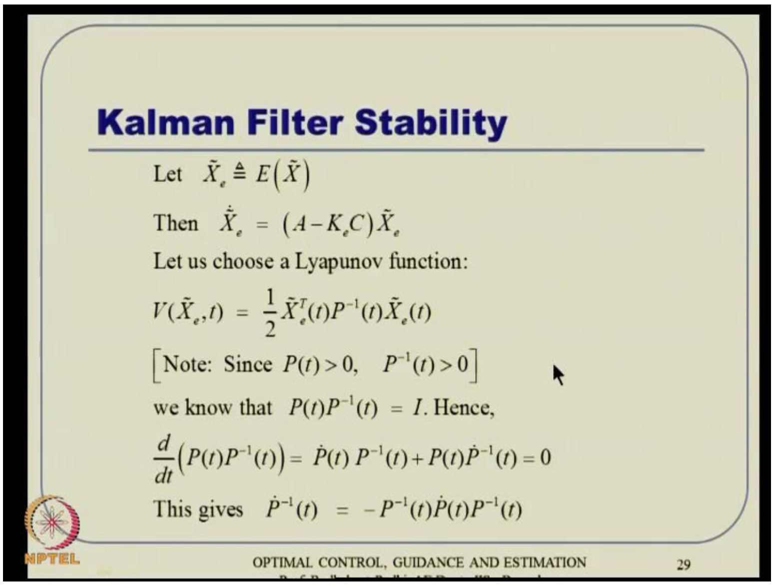
So, put it like here, and now you can expand this as a this buckets and excite this fact K e is nothing but P c transpose r inverse actually, and also we know that P dot is in this from, so if you expand that put this P dot expression expand and this is an algebra will turn out that v dot is nothing but this expression actually. So, clearly if it Q is positive semi definite and R is positive definite, remember, Q is positive semi definite represent hence, so this quantity this is symmetric term again and P is again positive definite term. So, this term remains positive semi definite, whereas this term R is positive R inverse is also positive definite and this is also a symmetric terms it transpose a sort of thing, so this turns out to be positive definite actually.

So, this means if the selection is made Q is positive semi definite and R is positive definite then V dot is guaranteed to be negative definite. So, stability condition is met, but this is local asymptotic stability, but also turn out tends out that this is the way we selected V is also radically unwanted actually, and it is also shown that it is actually a discrete time function basically.

There are all concept from non-linear control theory basically those of you interested can see some books that or you can see another course where it talked little bit on that on Lyapunov theory, they are some popular lecture, these you can some of those lectures to see some of this concepts actually. So, if the condition turns out that if v dot is negative definite, and an top up that if V is radically unbounded and it is also discrete time, than it is actually is s is satisfies this global stability behaviour and hence you can playing that this particular X tilde the dynamics is globally, uniformly, asymptotically stable basically.

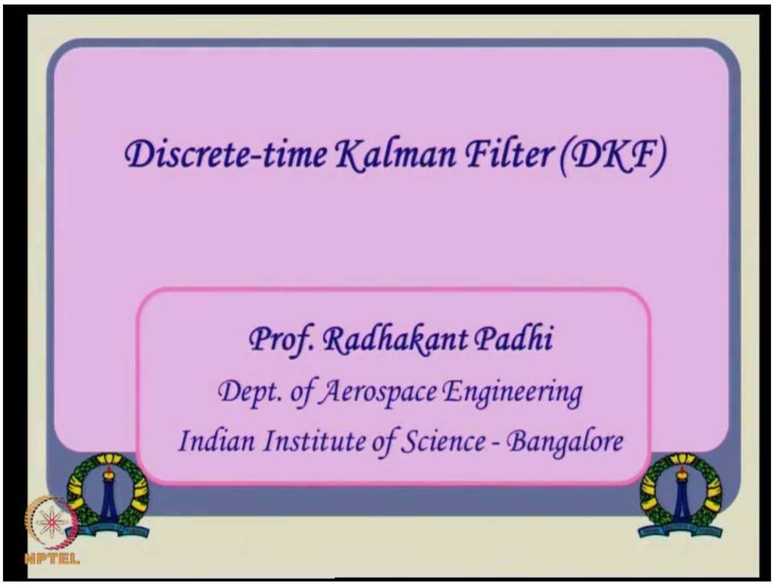
So, this very strong (()) sort of thing, strong result basically, so it is asymptotically stable and it is uniformly asymptotically stable; that means, it does not matter when you start symmetry l of t not actually. And it is also this result is globally true, that is what you are telling, globally uniformly asymptotically stable.

(Refer Slide Time: 26:49)



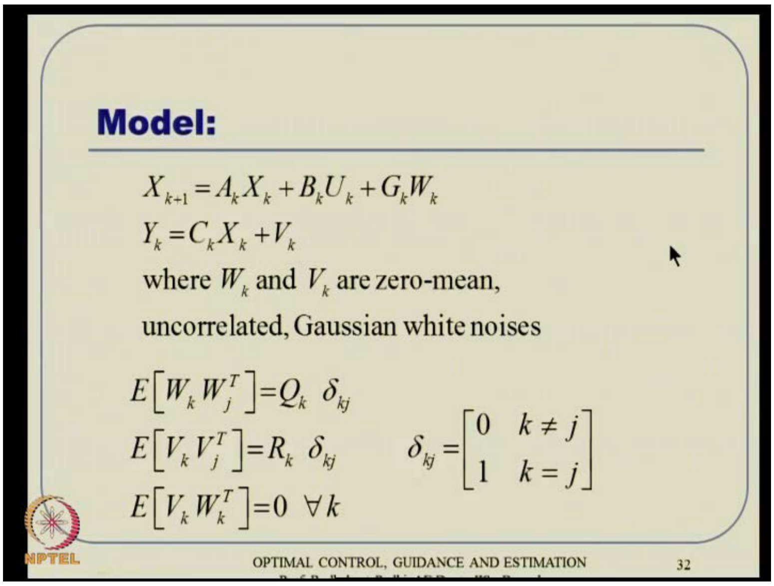
Alright, so this gives us a lot of confidence that nothing will go bad if we implement this, then will start arbitrary P naught it will take little more time to converge but it will converge actually, that is the generalization of kalman filter in continuous time from a from linear time in variant systems to time varying system, but I was talking sometime back will time back, but we are interested is a frame work where you can actually incorporate discrete measurement equation actually.

(Refer Slide Time: 27:26)



So, that leads to this concept of discrete time kalman filter; in other words, the system dynamics can be continues, and it is not typically continues, but you can mathematical discrete whereas, as a some sort of tools are typically discrete and we cannot make it continuous - I mean - it may not be physical because the sampling rate may not be sufficiently fast to interpreting that way actually. So, let us go, will see how you can derive everything in the form of discrete time and hence it does not max us too much on time invariant or time varying and things like that both will converge to the something, so in general we talk about time varying system dynamic - system discrete time moment.

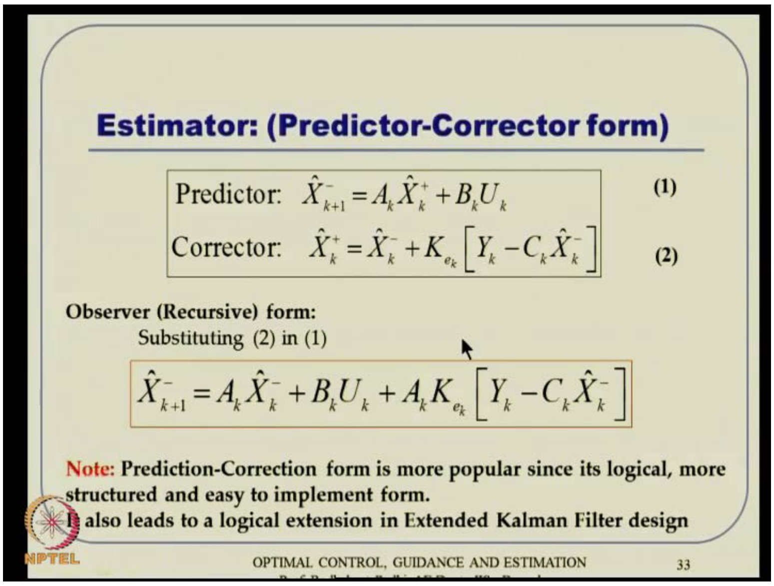
(Refer Slide Time: 27:56)



So, we have this system dynamics where X k plus 1 is given something like this, and Y k is something like this, k get stands for time step actually. And here this W k and V k again assume to be a 0 mean quite uncorrelated gaussian and white noise actually. So, what you are telling here is again the similar concept, but in the form of cronicre delta, (()) and this direct delta and cronicre delta kind of similar property, but they are - I mean - this operates some of the history domain and other one operates in the continuous domain actually.

So, that a cronicre delta K e define something like this if k is not equal to j it is 0, it is k equal to j it 1 which one actually. So, where you take expected value of W k times W j transpose there is some value which is dou Q k provided j is equal to k, otherwise if 0; similarly, expected value of V k times V j transpose is R k provided j is equal to k otherwise is 0, basically and become of orthogonal random thing like that are their no uncorrected and thing so; that means, expected value of V k ties W k transpose which turns out be 0 basically.

(Refer Slide Time: 29:36)

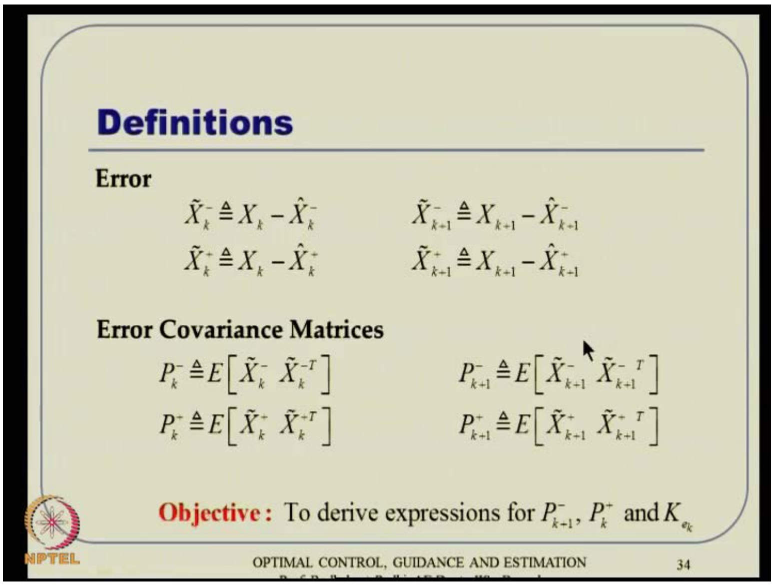


So, there are two ways looking at it and the first way is very popular and logical. So, one thing - I mean - this form tells I will propose it this way; that means, there is predictor equation and there is a correction equation corrected equation and if you can a substitute this equation two in one you can always get this form which looks very close to what you are done in continuous times actually. Alright, so this is actually observer form or recursive form what you called and, but this form it turns out to be much more intuitive and easy to implement and logically it means lot of sense also basically.

Here, e k what is the difference here is, here is all minus here minus, here minus, here and here if you some super state be minus some time plus, sometime like that and the implication is also - I mean - there, let me therefore, that implication first and come back to actually.

So, this do discrete time form, so the implication I was talking about implication of this. So, what it tends out this prediction you I think about something like staring at some sort of a plus value and going to this minus value using this correct - I mean - prediction sort of thing and starting with the minus value you go to the corrected value - plus value - using this measurement equation which is remember Y k is measurement at value basically.

(Refer Slide Time: 30:30)



So, what it means a something like I have this pictorial representation of something like this, if you see this time domain which is some k is transforming, this is k, this is k plus one, let say, which is k plus two something like this, and what happens is, I some value, I will predict, let say, this is X k plus I will predict, and I will land up with R k plus 1 minus. Then, I update this and once I update, i value will change to some other value, and that really I will call as r k plus 1 plus. Then I will predict again, and I will land up something n k plus 2 minus and here again there is a measurement, measurement can be either be positive and negative doesn't matter, finally, I will get n k plus 2 plus, so this way it will continue. So, in another words there is prediction on (()) I know some value that is prediction, correction; prediction, correction like that actually.

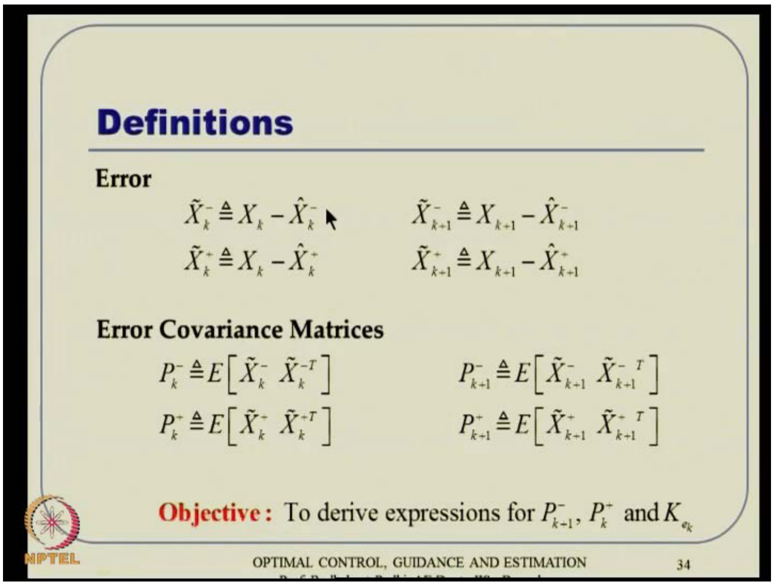
Alright, so starting with the updated value I can go for a predicted value then updated again and thing like that is why it made very logical actually and also make convenient for computer programming as well alright. So, that is this why written here prediction correction form is more popular since its logical, more structured, and easy to implement as well, it also leads to the logical extension into this E k for extended kalman filter when you see, it actually, this is the form that is most widely used and that is in E k f to mean actually prediction and then correction.

And also it gives us a platform, so just if you see this form, it gives us a platform that during prediction we do not really have to use this discretized equation, you can use another discretized equation, if you really want with higher accuracy actually, even though the theory assumes that you do it one step, one step all the thing, in the prediction stage from one can think about implementing some sort of higher order numerical integrations scheme that also gives a platform to implement that way.

Anyways, so this is what it is, so now the question is how we make sure the estimation is correct, our estimation is good actually, other words error of estimation is small, so that is what our objective all the time. So, when you talk about error now error is in enforce domain, think first of all he have he had something like, if you see this picture as well there was some value from which it was updated, so this is X k minus basically.

So, this is - I mean - if you see in this picture turns out that k, if I talk about time step t k or k, I have a value for X k minus and X k plus similarly at k plus 1, I have a value for X k plus 1 minus and X k plus 1 plus. So, essentially, if I look at this k and k plus 1, I essentially observe that there can be four errors actually, error can come in here, error error can come in there, error can come in here, or error can come in there, so then I have to define four error quantities.

(Refer Slide Time: 34:34)

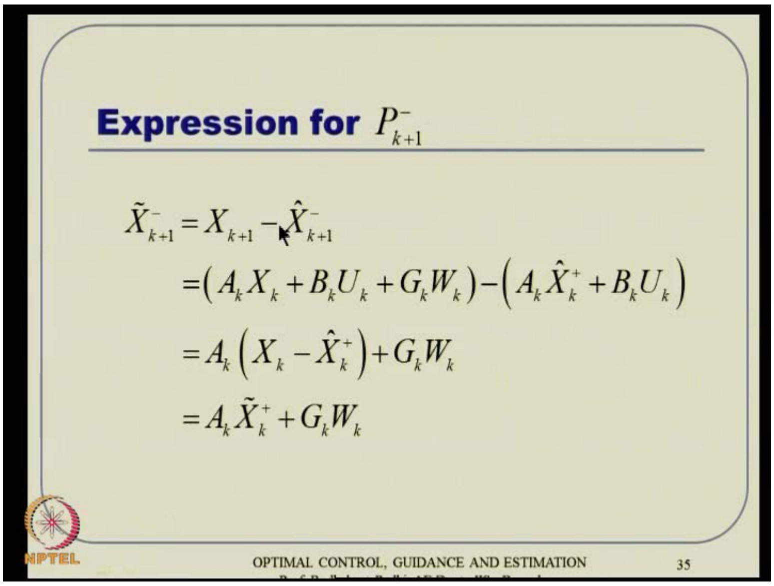


And first error quantity is X k tilde minus which is X k minus X hat k minus and similarly X tilde k plus which is X k minus X k hat plus.

Similarly, you can define that at time step k plus 1, now we define what is this error covariance matrix, and this error covariance matrix turns out to be something depending on this quantities, we have to define the corresponding quantities; that means, if I start with this quantity then error covariance matrix P k minus, P k minus turns out to be expected value of k, this quantity what about I have here X k tilde minus x tilde minus transpose same quantity. Similarly, if we this, if I tell p k plus, then I to take this quantity times this same quantity transpose, that is what written here similarly, things are here, and do not get confused too more with that much, it is all about good book keeping actually.

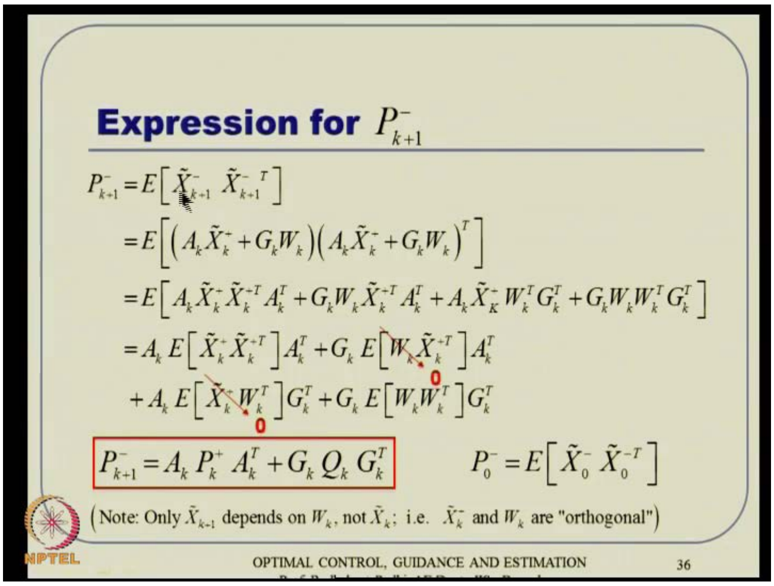
So, essentially what we are telling is we want to derive these expressions, and finally, select these kalman gains in such a way remember it comes here, in the correction equation actually. I want to derive this k e k in such a way that ultimately when I update my P k plus will turn out to be minimum, that is the whole idea there actually because when I update, I want to see a good update value updated value, should be very good whereas this quantity this P k plus has to be minimum actually. Similarly, p k 1 plus also needs to be minimum things like that actually. Alright to have these quantities different quantities we need to analyze this quantity first, because this is coming here and this one is related to each of the term here is related to the system dynamics.

(Refer Slide Time: 36:39)



So, we analyze this quantity error is nothing but true value minus the estimated value at k plus 1. So, true value is system dynamics, it comes from something like this, estimated value is something like this, from the equation - I mean - from the predictor corrector equation this quantity available b by k. So, once you put it back this k can be common, so you have this X k minus X k hat plus nothing but tilde plus, this quantity and then B U k and B U k will get cancelled out, this quantity gets cancelled out actually. So, we combine this first term with first term and leave the other one actually. Then, g k w k will turn out to be like this, so this expression is available now, so what about this, this expression is now available, so we can always talk about this quantity is expected value of this quantity times the same quantity transpose. So, now, we have got an expression for that quantity plus the same quantity transpose.

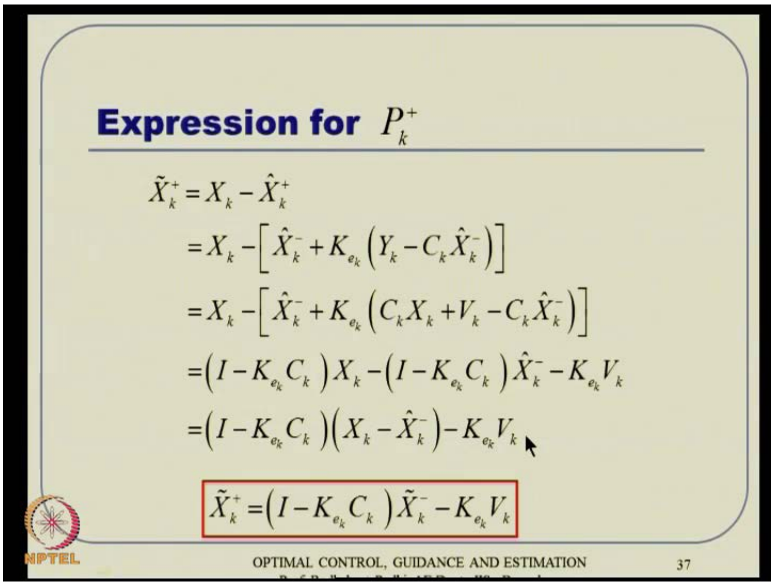
(Refer Slide Time: 37:28)



So, what you do now is expand this transpose, and then multiply this matrix, these two matrices, so we get four terms actually. So, if you do this algebra carefully, it will end up with something like this, and again X hat this fact that expected value is a linear operator. So, if we can separate it out all the things, but here you can observe that this there is a multiplication of W k and X k tilde plus. Similarly, this X k tilde plus and W k again so; that means, these are not correlated. So, essentially these two quantities will go to 0, so we are left out with this one and that one, this one turns out to be nothing but this, by definition and this one what we have here is nothing but Q k again by definition.

So, essentially this P k plus 1 minus which is essentially an estimate of how much the error it is after prediction turns out to be like this. And it obviously start with this, because, remember this is kind of a propagation equation, so we need some sort of P not minus value and it comes from there actually.

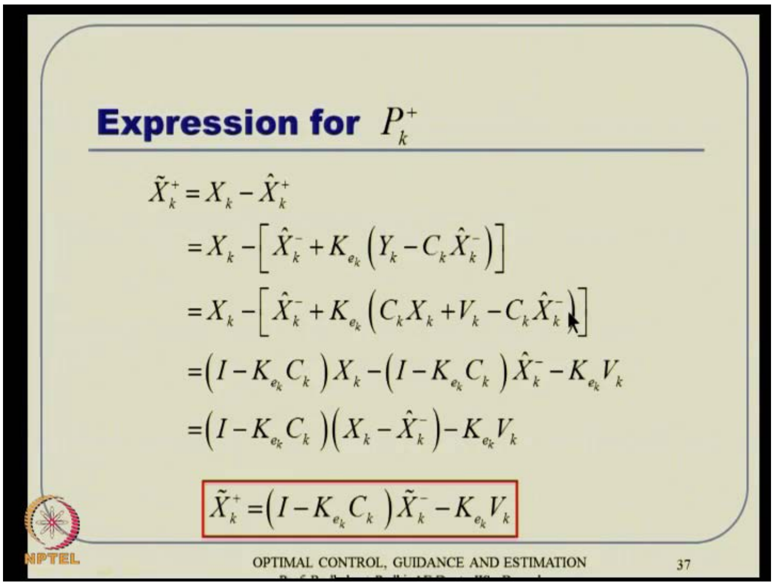
(Refer Side Time: 39:05)



Now, expression for P k plus then what is P k plus, essentially, expected value of this times the same quantity transpose, so we have to have some idea about this quantity, this is again by definition we go back to the definition and put something like this, X k is X k we written it like that but X k plus is this quantity, we have this X k plus is update quantity your correction corrector equation put it like their.

Now, we expand all these X k plus and now Y k is nothing but C k X k plus V k, so we put it Y k expression here and then expand all that, and it turns out that we can write it in this form actually. So, now here is k here is X k hat, so when you talk about this quantity, this is nothing but that by definition, they error between these two. So, error after update is a function of error before update which is very logical actually and then this V k quantity also. So, alright, so this is the type of thing so, that means, once you have - I mean - what you have now X k tilde plus which is something like this.

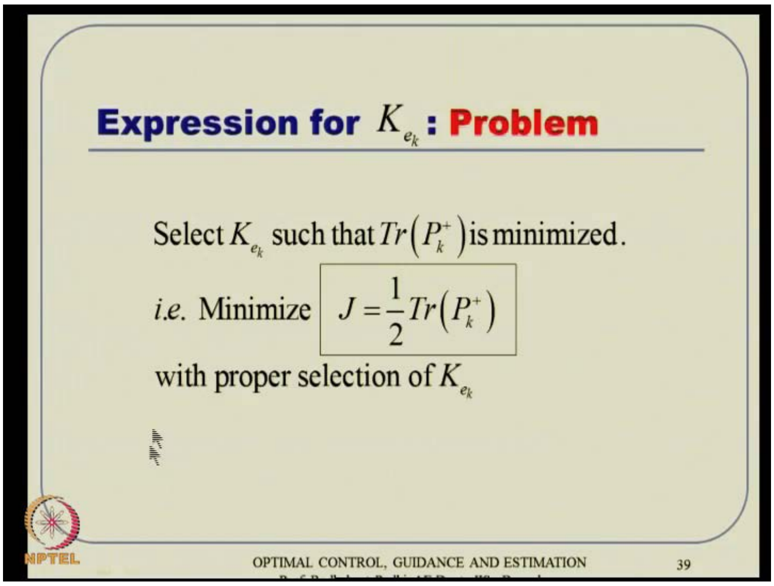
(Refer Slide Time: 40:36)



So, what we our aim is to analyze this quantity P k plus which is given as something like X k plus times X tilde k X k tilde plus transpose, so this quantity is something like this that is what we derived now, and hence we put it back here this quantity plus this same quantity transpose again, and we carry out this standard algebra and see - I mean - I suggest that you take a sheet of pen and paper and try to derive it yourself, then only you can see what is going on here very clearly.

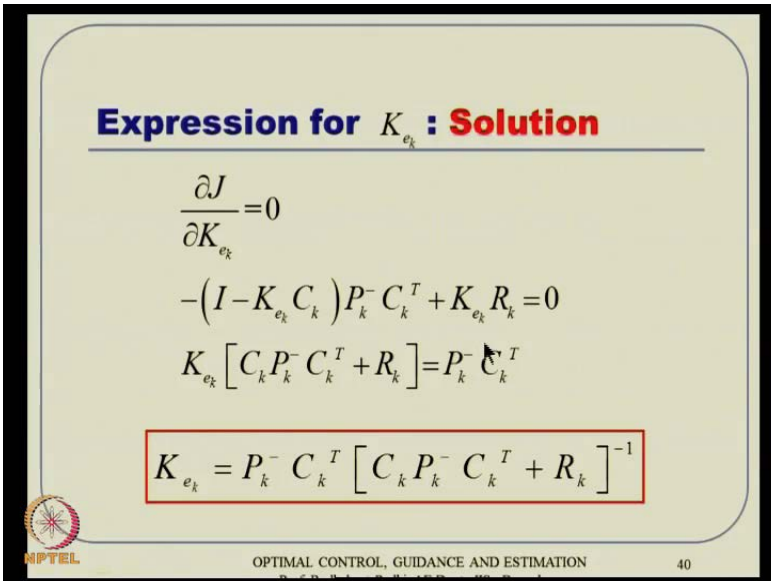
Then, ultimately turns out that we had this V k x tilde thing like that, that is not there, so that will go 0, so we will lend up with this quantity which is similar quantity that is nothing but p k minus. So, this term will be retain, this will go to 0, this will go to 0, and this term will be retain in the form of this, so we got a expression for error covariance matrix update actually.

(Refer Slide Time: 41:35)



So, now, what is our aim? Our aim is to have this quantity as small as possible, so after update at a particular time instant I want that estimation to be good; that means, the error between estimated value and true value should be as small as possible, and this is a indicator that so; that means, whatever P k plus expression is there I want to minimize that actually. So, essentially the problem is like this, we want to select a k e k in such a manner that trace of this P k plus is minimize; in other words, minimize j is a performance index half of trace of P k plus was a proper selection of P e k.

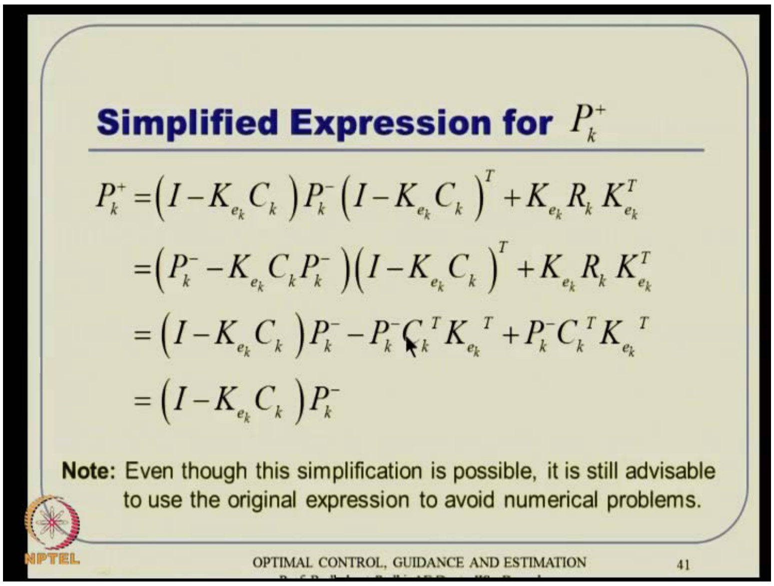
(Refer Slide Time: 42:17)



So, obviously, we have to excite this necessary condition that del j by del K e k equal to 0, so that del j by del K e k (()) this expression turns out to be like this, because j is nothing but trace of P k plus and P k plus is a variable here. So, (( )) using standard matrix algebra calculus - I mean - calculus for matrix expression we can derive something like this, this del j e by del K e k turns out to be like this. And here is actually a linear equation in terms of K unfortunately K happens to be in the left hand side in both the expression, so it is easy to solve.

If you take K e k in the left hand side take this, and then this P k minus times C k transpose happens with a minus sign, goes to the right hand side appears here and then K e K is ultimately something like remember, this is s right side product, so we have to multiply with right side inverse actually essentially what it means this K e k is P k minus times C k transpose which is this quantity plus this matrix inverse actually. So, what is holding here, we got an expression for P k e k, alright now P k minus - I mean - P k plus we got expression for that also.

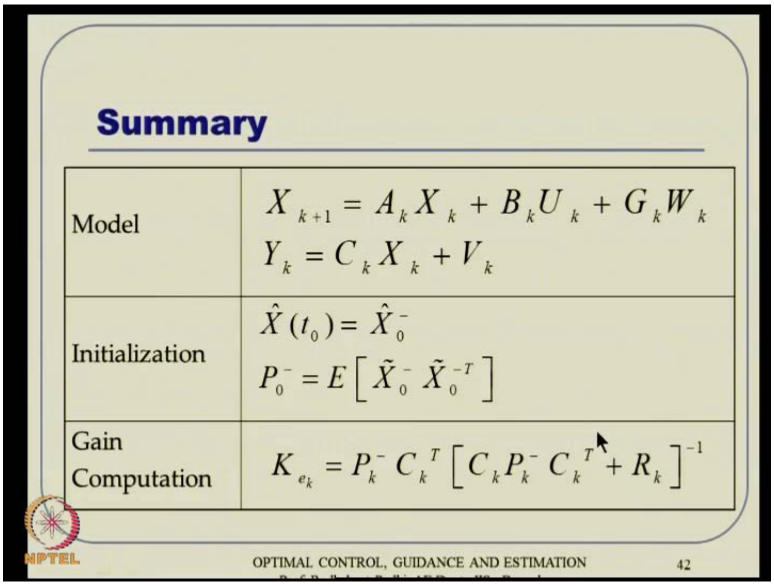
(Refer Slide Time: 43:44)



Now, P k because we remember, we need expression for update of this covariance matrix as well, so we got that and then we got an expression kalman gain also basically. So, we all kind of done, but a small interesting observation here that if we use this expression again, and revisit this expression then you can actually get it a little bit simplified expression, because we put it back - I mean - whatever we know here, if you put it back then, sorry, not that what I mean is, you start with this expression and try to simplify this. This transpose goes inside and then try to kind of expand this locate and thing like that, it turns out that these two quantities cancel out - these two quantities will get cancel out - you left out with that.

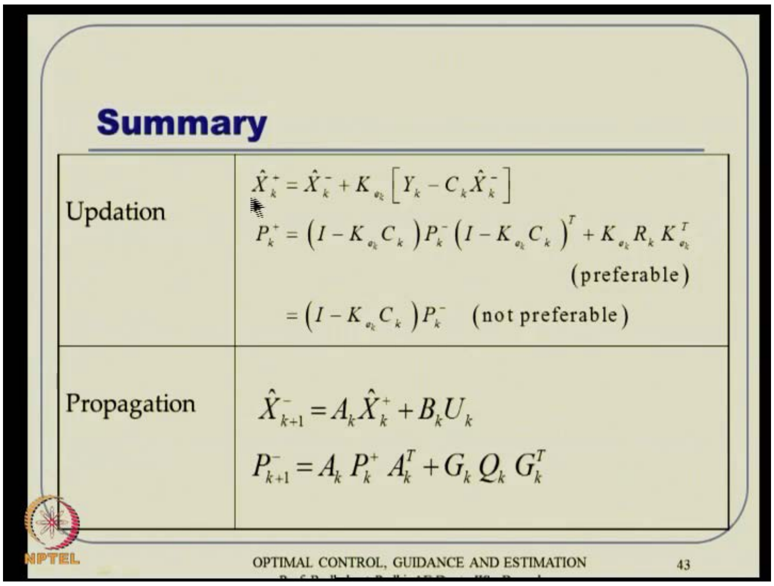
So, it turns out to be a most simplified expression, but unfortunately it is not a very good idea to implement this because, it will have numerical problems actually. This is not a symmetric expression, so in other words, you may this is chance that this symmetricity will be lost because of some other problems like round of errors and all that actually. So, there is a strong recommendation in the literature and books that even though this simplification is possible never get tempted towards using this, you still use this equation only.

(Refer Slide Time: 45:19)



Alright, so what is the summary here, after this we have this state equation in the form of discrete time, but discrete time varying system, and measurement equation is given some think like this, with usual assumptions that this W and V are white noise and correlated and thing like that. We have got, we have to initialize the filter and we initialize that way with values for X hat not minus and p not minus. Then we compute this which is all function of minus remember that, K e k happens to be function of R minus K values. So, evaluate that whatsoever, we have to update it actually, using this K e k we update this state equation in this form and here is that measurement will come and help us.

(Refer Slide Time: 46:05)

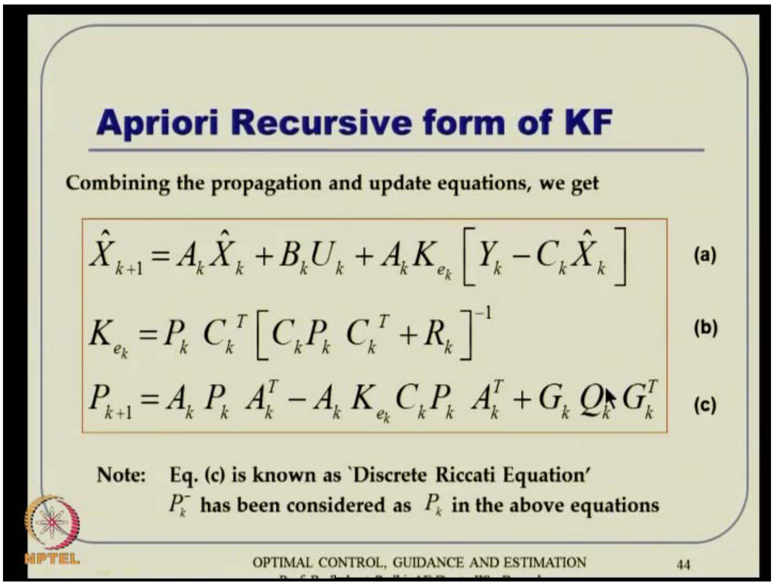


So, this is the kalman gain computational part, then we will use that for update, and we update both, we update state equation - I mean - state values and update the noise, sorry, the covariance matrix also basically - error covariance matrix. So, this is initialization, this is gain computation error state, and then this is update equation, we update the states like this, and we update the covariance matrix also like this, and again this expression is preferable, this is not preferable.

Then, we have this propagation after that the propagation turns out to be like this, this is system dynamics, we can propagate directly, and remember there is no noise which is taken into account, while propagating you cannot have a noise; noise is something that is unknown actually. So, we just propagate with the known part of the system dynamics, and then we propagate this P matrix also and P k minus 1 plus as P k plus 1 minus has been derived also basically - this one.

Using this expression we can propagate the P k plus, sorry, you can get an expression for P k plus 1 minus actually. So, this is update then you propagate then again go back to kalman gain by you compute that quickly and then update and then propagate like that actually. So, that will continue that way, so this is all about discrete time equation, implementation of kalman filter, but some people can also think we want to implement in a direct way, in other words, we can also go back and implement the direct recursive form actually.

(Refer Slide Time: 47:53)



And there is no point in having minus plus and all that there, because everything happens at the same time sort of thing, so minus plus superscript notation is dropped here, and you can go back until this is my - I mean - this observe equation where I need a K e K, so K e k I will compute it that way, where P k we need a P k, but P k will turn out to be, from this propagation we will obtain P k, we initialize P k, but using this equation, we can propagate P k actually.

So, this is the different alternate form, but - I mean - to my knowledge and many people will also prefer that the prediction correction form actually is very intuitive and easy for programming thing like that actually. Alright, so this particular lecture is good enough for understanding this, what you discussed here is continues time linear system, and revisited all that, and then using those derivation ideas we had derived these same things for time varying linear system as well, then we went to this the discrete time form and then we have this derivations of all these equations, where we have an idea of how to implement it, both in prediction, correction form as well as this direct form actually.

Alright, so in next class onwards, we will go back to the real problem of non-linear systems, and try to see what way we can extend these ideas for these external kalman filter and beyond actually which is, what is used in practice actually, alright so this much in this class; thank you.