Optimal Control Guidance and Estimation

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

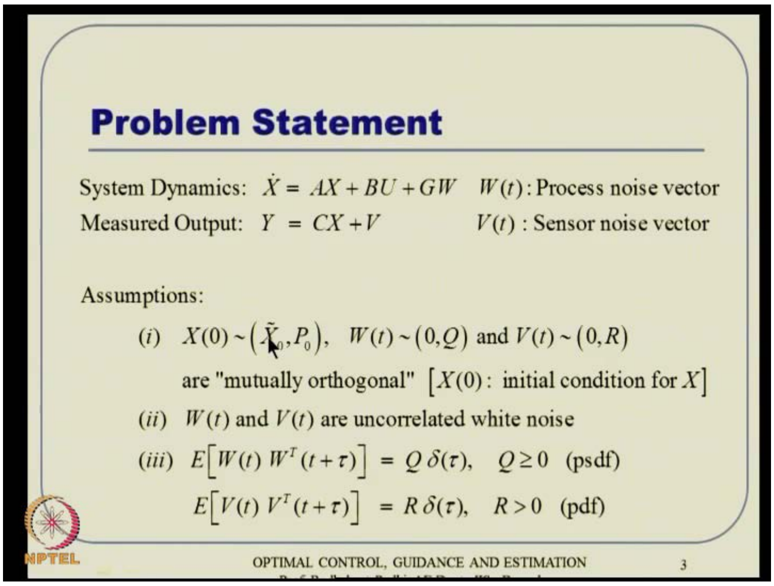
Lecture No. # 28

Kalman Filter Design - 1

Hello everyone. We will continue our lecture series with Kalman filter. In last couple of lectures, we had given some overview and some implementation formulas and things like that. How do you use Kalman filter design and also I told in last lecture that we will actually derive some of these relationships to have better understanding around that way, ok.

So, this is where we will start our derivation process and first thing is we will derive everything in the linear domain and continuous time domain actually. So, let us understand the theory behind that. So, first thing what we are interested in this particular lecture, this Kalman filter design for linear time invariant systems in continuous time domain, ok.

(Refer Slide Time: 01:03)



So, going back, I mean the problem statement that turns out to be something very precise like this. We got a system dynamics which is X dot AX plus BU plus GW and where we measured output is see something like CX plus V, where W and V are noise things and W is the process noise vector and V is sensor noise, I mean sensor noise vector basically.

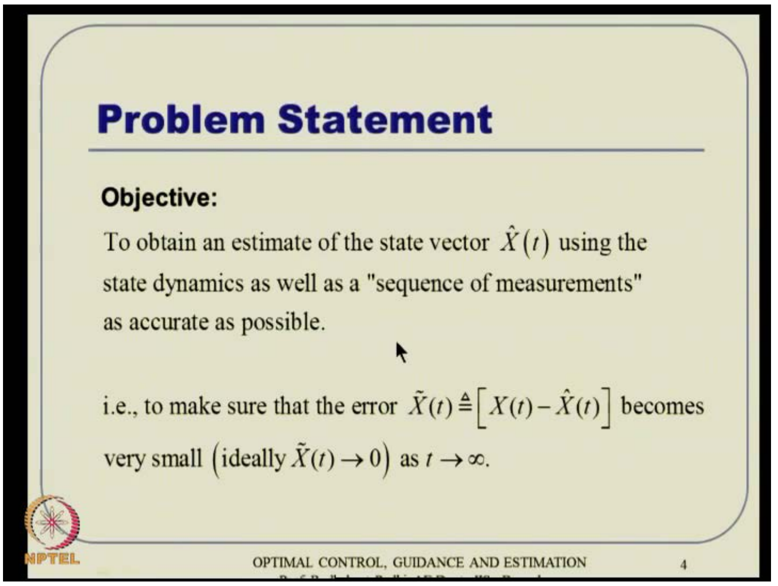
So, there are bunch of assumptions which will make our life much easier later and since, it tends out to be something like this, ok. When anytime you have the linear system dynamics and remember these are like time invariant system. That means, a b c d all that are constant matrix actually and then, anytime like dynamic equation, we also have initial condition associated with that, ok.

So, the initial condition, that is X of 0 is assumed to be this way X tilde 0, P 0 that this notation till this. The first element is the expected value mu or mean value and this is nothing, but covariance matrix actually, ok. So, X of 0 has a non-zero of mu. Obviously, that is where you expect things to be happening in the beginning and it also has something like noise covariance way like this error covariance matrix P naught actually.

What about W and V? W is w and V expected to be I mean they are characterized as 0 mean. So, both of them are 0 mean. W h Q is something like process noise covariance. So, that is how it is. Q is defined and V, let the covariance is R actually, ok. What you mean by that? Obviously, mean something like this. That means, expected value of W. W transpose if you take that way, it turns out to be yet Q times the tilde function and similarly, expected value of e V transpose if you take, that turns out to be at A of tilde function multiplied by R actually, ok, but there is a very important behavior here that the two assumptions actually which tells us that X of 0 W t and V t are actually mutually orthogonal. That mean I take any common X of 0 and W r W to V or V to X 0. At any combination, they are mutually orthogonal, ok.

Second thing is this W and V are non-correlated, uncorrelated and they are white noise. So, see the assumption. Actually many assumptions are involved, but V to zither filter still works actually, ok. As to summarize again, here I have got three things refers randomly, varying and initial condition W and V. They are characterized by something like the mean value and covariance matrix everywhere. Can you make sure that the X hat and true X goes to 0 basically? Then, we will get X hat which closely resembles X actually.

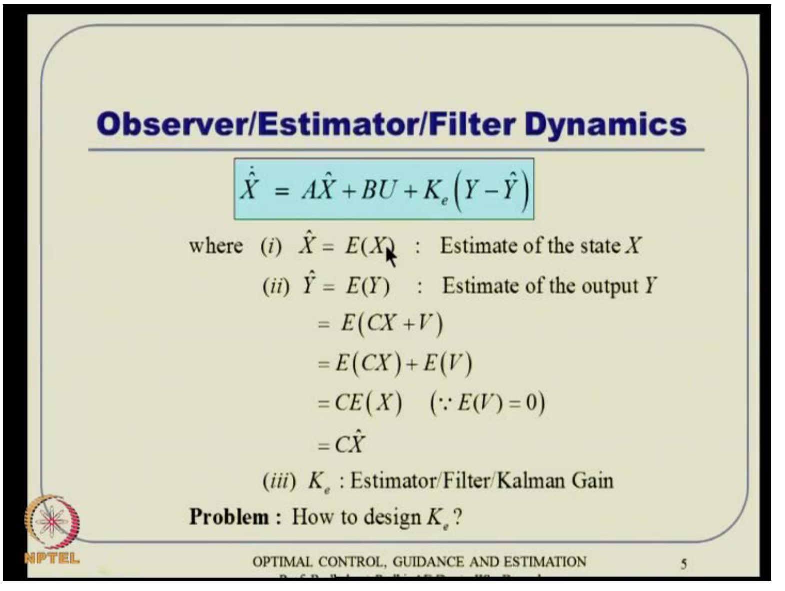
(Refer Slide Time: 04:25)



So, this is the process that we are interested in getting some exact of t and in the process, we will use state dynamics as well as a sequence of measurements. Then, you will have some estimate X hat of t in the sense that the error between X and X hat, if I define X tilde as hat, then it goes 0 as t goes to infinity, ok.

Initially, they will have some error, but V whirls this error will not be there basically, that is mean X hat will closely weak or closely resemble X actually. What is helping us in doing those two things? One is as the system dynamics and the other one is the estimated dynamics. So, the other one is measured output, ok. All right. So, let us see how it is possible. First thing is I mean proportion observer dynamics or sometimes estimated dynamics or filter dynamics. People can say it in different names and all that way.

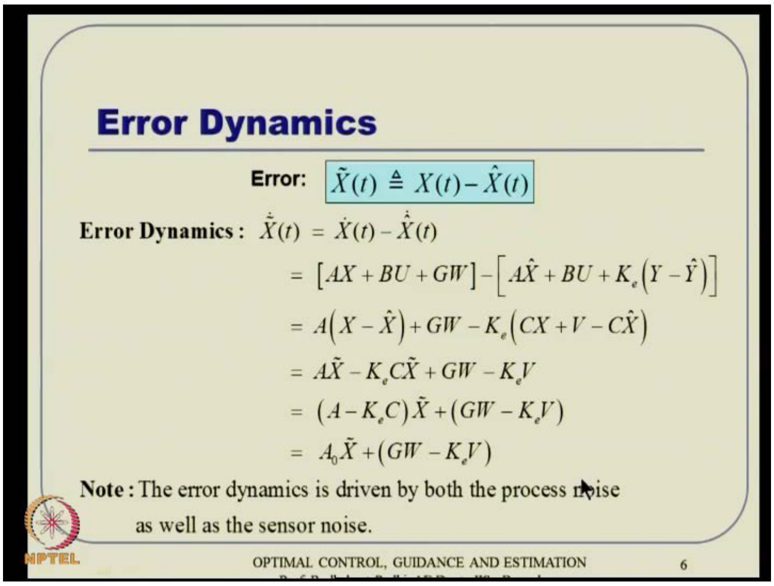
(Refer Slide Time: 05:59)



So, this is defined very close to what we already have here. X hat dot is the AX hat plus BU plus K e time innovation Y minus Y hat. Now, Y hat is an expected value of Y basically, ok and expected y is nothing, but CX plus V. So, you can substitute that CX plus V, but again expected value is a linear operator. So, you can do this linear operation here expected value of CX plus expected of E, but expected value of V is 0. That is 0 mean white noise actually mean is 0. It is gone. Then, its again expected value is linear operator that turns to be C times AX and E of X is X hat actually.

So, K e, what is happening here is nothing, but the gain actually. Estimated filter or Kalman gain and the whole point is how to regain this K e. Second you know this K e, we have got this initial condition at least the mean value sense and we can strictly propagate these dynamics. That is the whole idea here. What you have to design K e in such a way that this happens that may error goes to 0 as t goes to infinity. This is the primary objective actually. How to design that? So, we define the error X tilde is X minus X hat and the error.

(Refer Slide Time: 07:36)

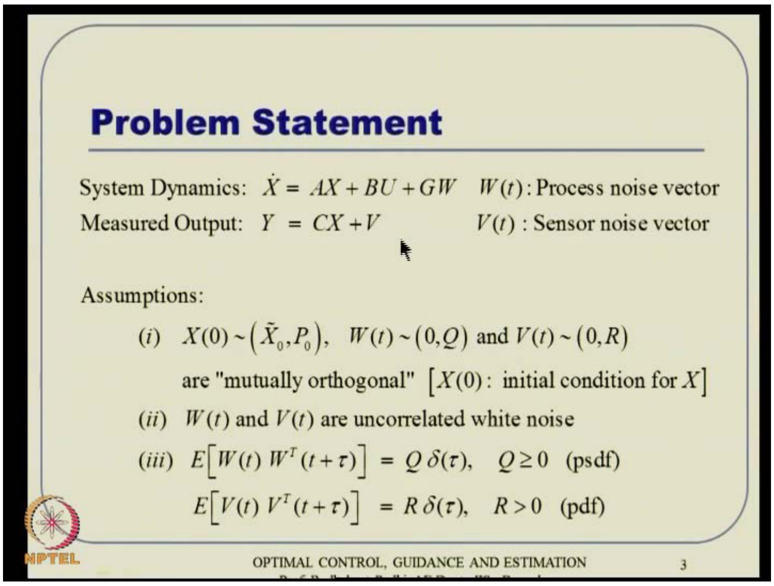


Then, we talk about something like an error dynamics, how the error propagates with time, ok. So, X tilde dot is nothing, but X dot minus X hat dot simply from this definition and then, X dot is nothing. This follow AX plus BU plus GW, where X hat dot is nothing, but the observer dynamics and observer dynamics that you put that. Then, we combine terms. So, AX minus AX hat. So, that comes here, GW comes here, BU, well this BU VU gets cancelled out, ok.

Then, you get minus K e times, sorry GW, GW here. Then, minus K e times Y is nothing, but CX plus V minus Y hat is nothing, but CX hat. We just derived actually as of this CX hat. So, this term is nothing, but X tilde. So, this is A times X tilde here and then, minus K e times, I mean this. Well, let us this term this is the GW minus K e here, ok. Then, you have got these two terms. You can combine X minus X hat is X tilde. So, that is K e times C times X minus X tilde. So, that is XX minus X, that is X tilde basically.

So, here we got AX tilde coming from here, GW minus K e V goes there and whatever remains K e times C into X minus X hat which is X tilde comes here, ok. All right. So, then, we can combine these two. If X tilde is here, X tilde is here you take common ousted dot that becomes X tilde here and left out with that term actually, ok. This particular thing you can define it as something like as 0. Then, it turns out to be S tilde dot is nothing, but AX0 tilde plus these quantities GW minus K e of V, ok.

(Refer Slide Time: 10:12)



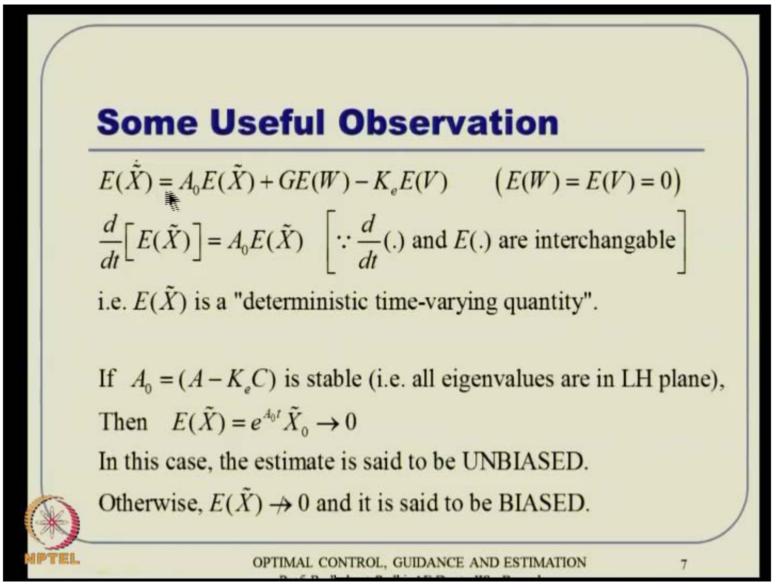
So, what happens here? The interesting observation here is the error dynamics is now driven by both process noise and sensor noise. The dynamics we call true dynamics is actually driven by process noise. What is the moment put here filter in the loop something like this and then, talk about error. Then, the error dynamics is nothing, but a function of both W and V. That means, it is a function of both process noise as well as sensor noise.

Now, some useful observation here that is if you talk about expected value of X tilde dot, then what happens is you take this expression and then, put expected value and then, expected the linear operator. So, we can all do that and it turns out that this G e is a constant matrix; K e is a constant matrix and all that. So, if you this GA constant matrix, K e is a constant matrix, it will come out expected value the linear operator again and then, this one and that one turns out to be 0. These two expected values of W and V 0 of the 0 mean actually.

So, 0 mean means expected value of them are 0. So, this turns out that the dynamics in the expected sense, expected value sense d by d t of expected of X tilde is nothing, but A naught times expected U, I mean sorry expected value of CX tilde, ok. That means e of X tilde. That means expected value of X tilde is a deterministic time varying quantity now. Remember that you find A minus K e C. So, that is deterministic respected value is a mean value that is the deterministic operator actually. I mean once you operate an expected value, the result turns out to be a mean value number, basically that value.

So, when you talk about the expected value of the error and then, its dynamics d by d t that this governs something like that. So, what it tells us an expected value of X tilde is a deterministic time varying quantity basically.

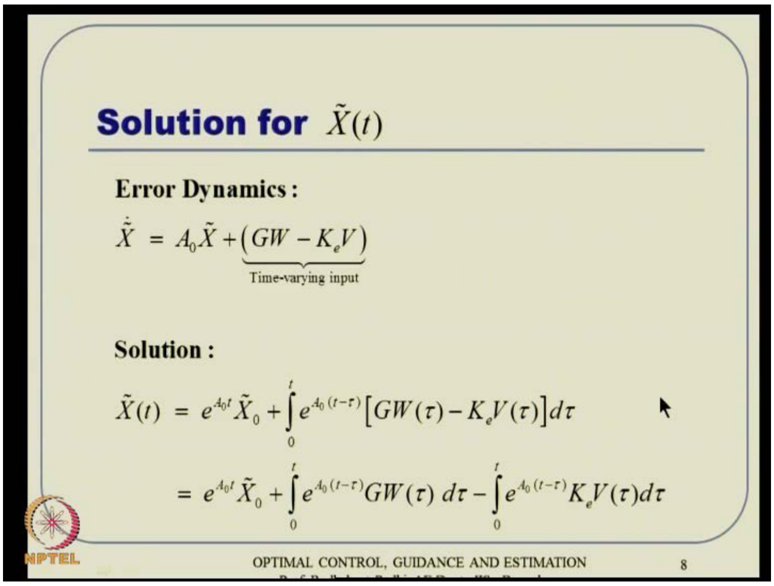
(Refer Slide Time: 10:21)



Now, it turns out that this naught which is defined as a minus K e C is stable. That means all Eigen values are in the left half plane. Then, what happens is expected value of X tilde is nothing, but we take a solution of that is nothing, but to the power A naught t times X matrix actually. X 0 tilde and that will get to 0 basically because this matrix is always stable. If it happened that way, then it is this said to be the estimate is said to be unbiased because the expected value ultimately goes to 0.

What if it does not happen? Now, that means, expected value of X tilde does not go to 0, then it is said to be biased actually, ok. So, you will ultimately result in some sort of a biased estimate which is not really good. This error thing should go to 0 basically. Then, you get what is the true valve in equations.

(Refer Slide Time: 13:05)



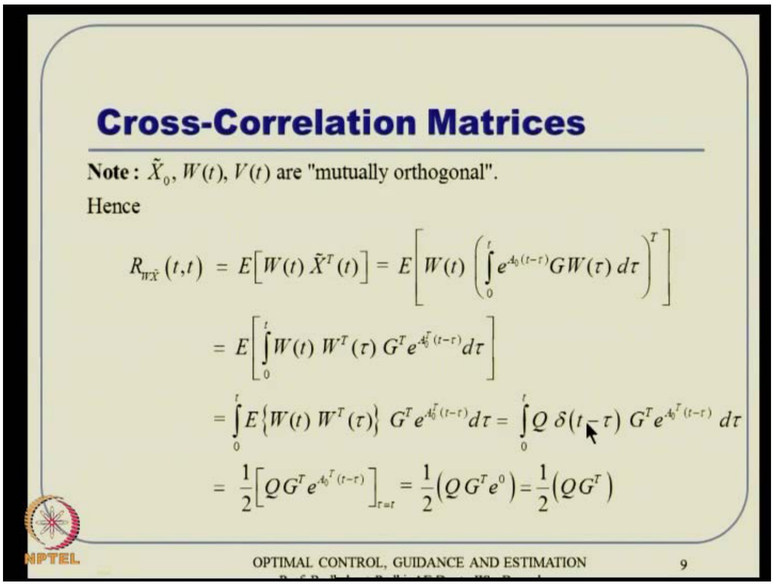
So, now, we will constitute this. Going back to this, we will construct this Kalman filter slowly and first thing to see that this aerodynamics is given something, given by something like this. Then, how you use it actually? So, we ultimately need some of these expressions. Later we will see that these expressions are P needed actually. So, we will go slowly actually that way. Then, first thing is what the solution of that is? The solution of for getting the solution you can think this part is nothing, but a time varying input is randomly varying, but this still a number basically. So, still it goes to the dynamics and then tries to alter it.

So, it is a time varying input. If you see that way and once you know this is the time varying input, the solution turns out to be like that from linear systems theory. So, this part is e to A naught times X tilde naught or homogenous part plus cost C function part which is given something like a convolution integrated basically e to the power A naught times t minus tour multiplied by all these actually and d tour, all right.

So, this is the solution of that considering this as something like a time varying input. So, what if you simplify this? Obviously, these two can be separated out first. It means this GW and K e V part of it and then, we are interested in this RWX tilde basically and here, we will use this property that these guys are mutually orthogonal. All the signals are mutually orthogonal. So, when you compute, when you attempt to compute RWX tilde, then this is nothing, but expected value of this follow and extend by you just it and that is the reason why you want it actually, why you want it by the way ok.

So, you put this X tilde expressions, all right. So, you put that because remember if I, even if I take this below, then using this linear properties of expected value and because this is orthogonal to each other and all that will be cancelled.

(Refer Slide Time: 14:32)



So, what will remain is only the integral part coming out of here and then, we will keep, I mean only that part actually really W because when you operate it with W, only W part will remain. Remember, W and V are also mutually orthogonal. That part will also go. So, this will go because of mutual orthogonal. This will also go because of mutual orthogonality. What will remain is only that part actually.

You put it this, that only that part and then, you can tell expected value of this is nothing, but transpose thing I can take inside now, but the sequence will be reversed and this I can attempt to put it inside because this is my integral 0 to t. So, W of t is not a function of tour really. So, this is like a constant s with respect to this integral. So, I can keep it inside. So, that is what is done here. W of t is taken inside. Then, expected value is a linear operate integral is also a linear operator. So, these will be coming out.

So, expected value can be taken inside and remember, these things are actually not random variable. They are deterministic. So, they do not have to be a part of expected value. So, expected value turns out to be this, but this by definition is nothing, but this Q times tilde delta function actually, right and if this is an integral evaluation with a derived delta function, there is interesting property actually with of any integral evaluated with derived delta functions. If you have A f of t and evaluated of delta, sorry T minus tour a sort of thing actually f of t derived delta of t minus tour.

Well, let me not do that way fine, thus then d tour sort of thing. If you well let me not do that, this is just tour. Let say will it say tour d tour and integral is there actually. So, what happens here is like this. This is also tour by the way. Well, in general, we can do this one. This is let talk about got confused here. In general, we can talk something like this, f of X. Then, delta of X this this direct derived function the x yet to be lets say this will turn over to be simply f of A or f of B. Let us say it is half of A. If X is equal to A e X equal to A. This is something like half of B if X equal to B, ok.

Well, I know it does not work and this will turn out to be simply f of I mean any value f of whatever that C sort of thing. If x is any value between X equal to C, where C belongs to AB, strictly AB not included. In other words, if the value turns out to be naught of the boundary value, then the integral value is just half of that half of the value at that particular point, half of the function value at that particular point and if it is strictly inside the interval, then it is just the function value basically, but what happens here, interestingly if you observe it, what happens here is the integral value that we are talking about is actually, where are we, here. So, here we landed up with and then, if we notice that when tour equal to t that kind of thing, then that happens to be a boundary value.

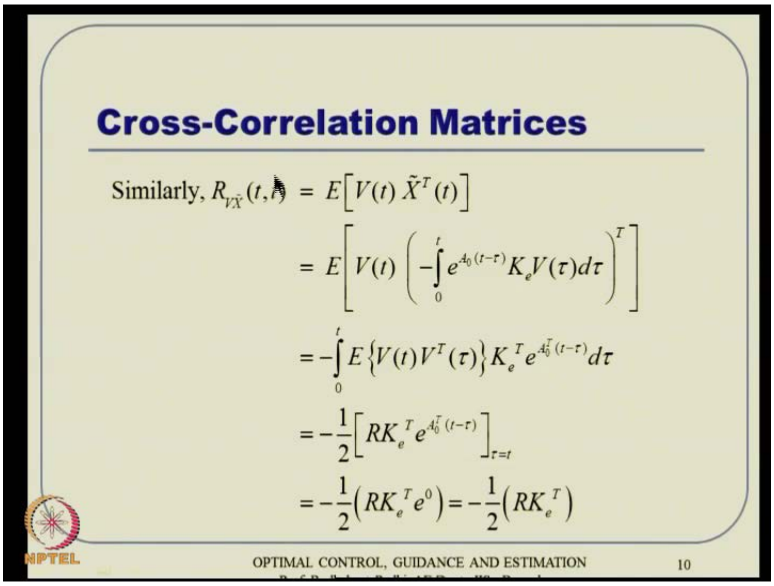
So, the entire function whatever it is, we are having we can simply evaluate it t, but we have to make sure that is a half terms X into that. So, the half QG transpose e to the power A 0 transpose and this is 0 basically. So, evaluate it tau equal to t basically. So, when we talk about this one, this is e to the power 0 and e to the power is identity. So, that means, we land up with some value like this.

I hope it is clear because this is one of the understandings that we need to have. So, we derive the solution like this and then, tell we are interested in this operator or WX tilde. Remember, W is in deep, I mean the W is mutually orthogonal with respect to X 0 tilde and R t. So, here is X to the power tilde and here is V tour and all that. So, these will ultimately go to 0.

So, we are interested in that what will not go to 0 is only this part. So, we keep that part only. Then, we use the idea that K expected value is a kind of a linear operator and this part is constant with respect to this integral because t happens to be one of the limits. This can go inside and then, subsequent expected value can go inside the integral. Then only, this expected value will turn out to be and that is nothing, but a derived tilde function times Q and this direct because this derived delta is there as integral actually and this integral goes to 0 at tour equal to t which is actually A. I mean it is getting evaluated at tour equal to t sort of thing. So, that is some, I mean this is the limit of the interval and this is because of that the half term will come and this is what it is and because it has to be evaluated at tour equal to t which is 0. That means, e to the power 0 is identity. So, we will land up with something like this actually, ok.

So, this is something that or if you are still not very compatible or something, it suggest that you see this derived delta function, integral evaluation and all that. It is available in any text books actually. So, on mathematics rest out of that actually. All right. So, this is our t actually.

(Refer Slide Time: 22:40)

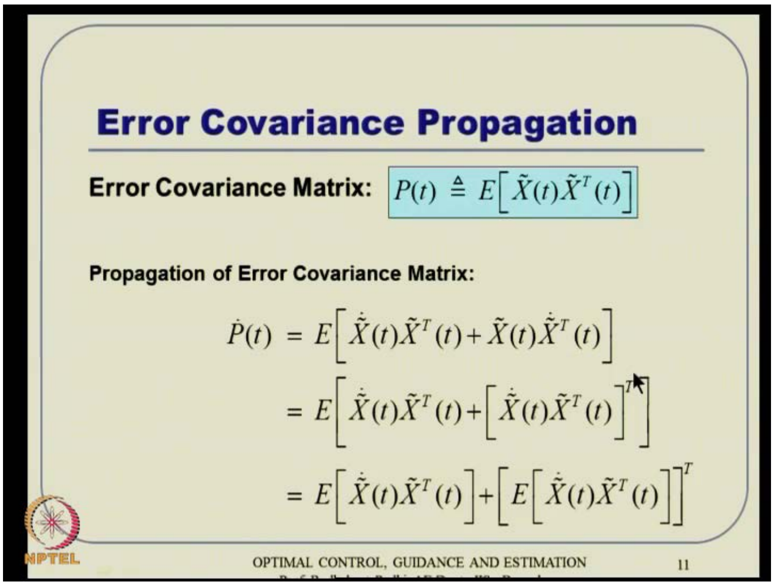


Similarly, now this is one term that we got. What about the other term? RVX tilde. So, this VX tilde is again the similar procedure. Now, we have to keep the V part of it because these two will go to 0. So, once you have V part of it and then, there is A transpose this or take the transpose inside side and then, reverse sequence will happen and again expected value and integral of a linear operator. So, one can go inside. You will have this W, w transpose again which is nothing, but Q. Again, there is an integral evaluation and that is going to this, goes to 0 at one of the boundary values because its half term will come.

So, again this turns out to be this below and sorry, the RV of X tilde will be very similar. You remember this is a negative term here. So, this negative term will be there. Here, that part again if we take V t inside and then, this transpose inside, it is a reverse sequence. So, then whatever its expected value will go inside the integral, then this part is nothing, but R times derived delta basically and then, again this is evaluated at one of the limit points. So, half of that and then, evaluate it tour equal to t as its identity here. So, it is left out to be that actually.

So, you are interested in so much of this analysis of these two things and all because we will soon need it in the error, I mean the covariance matrix propagation and things like that. You need it actually.

(Refer Slide Time: 24:01)



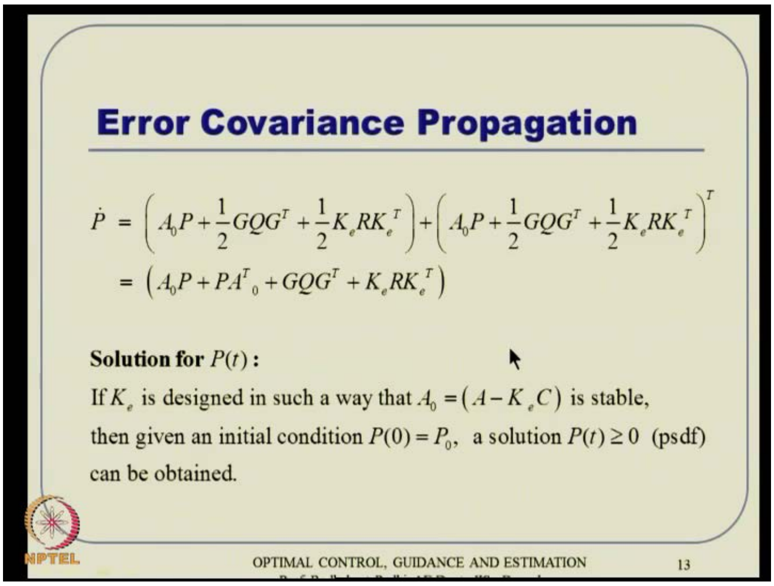
So, what is error covariance matrix? As I told, it is defined as an outward product between the errors. This variable X tilde, X tilde and X tilde transpose expected value, ok. So, what about the error dynamics? I mean this propagation of this covariance matrix what happens actually. So, then, that is P dot and remember derivative and expected value, both are optimal linear operator.

So, derivative can go inside the expected operator and I have two variables like this. So, it operates that way and remember, these are vector matrix things. So, make sure that the order is maintained and then, this part I will keep it as it is, but what about this. This is nothing, but I mean if I alter these two then, the whole transpose actually and I take respected value inside. That turns out that it is nothing, but a expected value of this plus expected value of this plus expected value of the same thing whole transpose actually, ok.

So, if I get expected value of this, I do not re-compute it because this is nothing, but this same thing with the transpose thing. Now, what about this below? So, expected value of X tilde dot X tilde transpose, what is that actually? So, now X tilde dot is nothing, but this one, right, we derived before. So, you just take this one a lot of simple book if in actually. So, then you multiply, expand the bracket, multiply A anywhere, then take an expected value and remember, this is what is expected value. What you are seeing here is nothing, but at WX tilde and this expected value to see nothing, but RVX. That is why we are interested in these.

Then, this by definition is simply P. So, what it turns out to be this one is nothing, but A naught P plus this quantity which we derived to be G G times this quantity and this quantity is derived like this. Why not K e times RVX tilde? In RVX tilde G I something like this, ok.

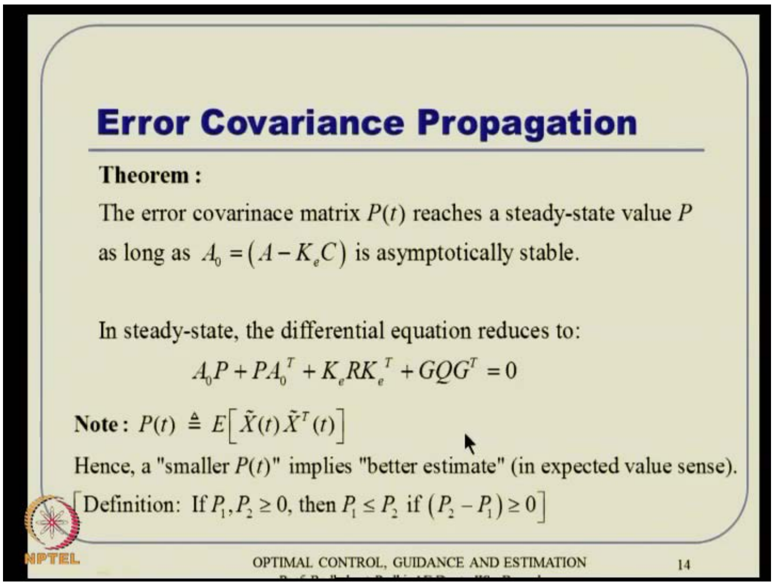
(Refer Slide Time: 26:29)



So, then if you expand that it turns out to be A naught B plus half GQG transpose plus this minus. Minus become half of A K e R K e transpose basically, ok. Then, what is P dot? Remember P dot is expected value of this plus same thing whole transpose. So, you got expected value of this. So, then, P dot is expected this one plus the entire thing, whole transpose actually, ok.

So, P dot terms out to be like this. So, very close to what we know in LQR theory actually, ok. All right. Now, we need to find the solutions of P and the theorem tells us to somewhat result is there which tells us that if K e is designed in such a way that A 0 or A naught which is defined as A minus K e. So, if this A minus K e is stable, K e is designed in such a way that A minus K e is stable. Then, even in an initial conditions of P which is P of 0 is P naught, then a positive, semi-positive, semi-definite solution can always be obtained. That is what the P times actually is, ok.

(Refer Slide Time: 27:29)

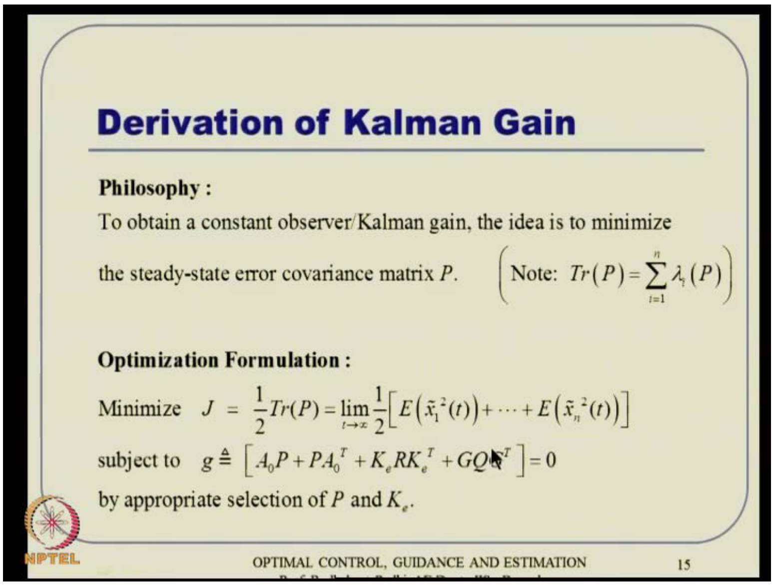


There is also naught theorem which tells us that the error covariance matrix P t actually approaches a steady state-value P as long as this is asymptotically stable actually. Not only there is a positive definite solution, but it actually goes to some sort of a steady state valuation actually. If a steady state solution and we want that particular steady state solution, only then P dot has to e 0 basically, ok.

So, this entire expression, this one is nothing, but P dot, but P dot is 0. So, hence the entire expression equal to 0 basically, all right. So, in steady state, the differential equation reduces to something like this, ok. The need to not I mean some comment here that P of t by definition is something like this and hence, what you mean a smaller P f t implies, but I estimate basically, ok.

The error covariance matrix is smaller once the estimate is better basically, ok and there are also some definitions like this which tells us that if P1, P2 or both positive semi-definite and P 2 minus P 1 is, I mean 1, then P 1 is less than equal to P 2, provided P 2 minus P 1is a positive definite matrix. Remember why this is this definition sort of thing? It is because we are not talking about the matrix algebra basically. So, you cannot compare these two. Whenever you have this P 1and P 2 matrix, how do you compare them and by definition, it tells out that if P 1 is less than equal to P 2, if P 2 minus P 1is a positive semi-definite matrix basically.

(Refer Slide Time: 29:20)



Now, here is a point where you can actually go ahead and try to derive Kalman filter. So, what it tells us now so far. So, actually if you see all this, our aim was to get something like AP dot and P dot. If you see a reverse sequence sort of thing, our actually aim was to get some sort of a error covariance matrix propagation dynamics actually. So, we wanted to get AP dot and P dot turns out to be something like this and the same value to whole transpose actually, ok.

So, then this value we are interested in, so we had to do some system only like that, but here is a client where you get this R double X tilde and RVX tilde. So, we wanted that expression actually and that is why, we derive this RVX tilde and RWX tilde that way, ok. So, now, if people like this P dot T actually, so we got this P naught and some theorems which tell out the system solution. If A naught is stable, then there are always existed P positive semi definite P and in general, I mean it increases the steady state value. Once it is as a steady state value, the equation is something like that actually.

Now, what you are interested to obtain a constant observer Kalman gain. We are not interested in time varying K e and all that. We are interested in something like a constant K e basically.

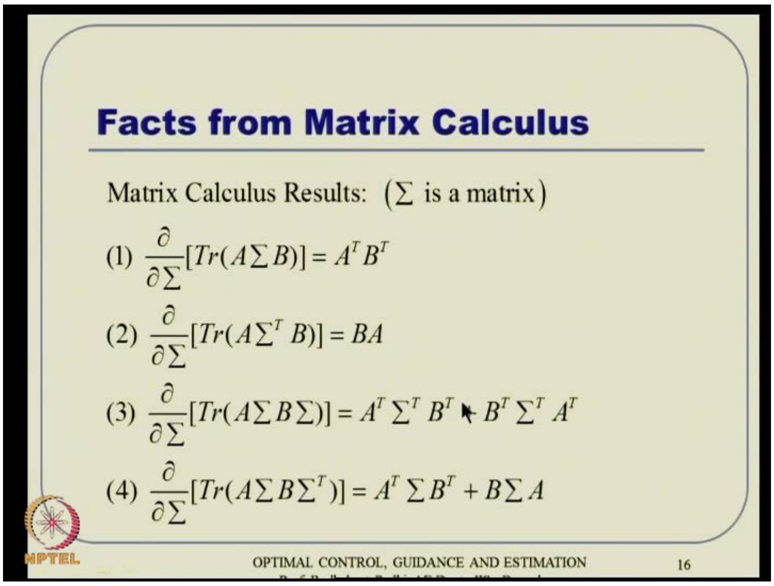
So, if you do that, then the idea here is to minimize the steady state error covariance matrix. If you look at this one, the linear solution which is positive definite and positive semi-definite at least and on that, ok, so that is what we are telling here were interested in a steady state value which is smaller actually, ok. That means we are interested in minimizing the steady state covariance matrix P basically. What is steady state covariance matrix P? It is the solution that comes out of this basic equation actually, ok.

So, here is an optimization formulation which tells that we have two kind of minimize this P. What? Remember P is a matrix now, but rest turns out to be a kind of a non-basically. So, we are interested in minimizing the trace of P known that is, ok. So, what is P? By definition limit, it tends out to be infinity. This term actually, right just the trace of P. What it means? Then, limit t tends to infinity that is what we want in this.

So, when t goes to infinity, this error quantities are there, X 1tilde square XN tilde square like that actually and expected value of all that. If you put it in some term in that is what you want to minimize actually, ok but this has to be minimized to subject to this constraint equation. This is what the constraint equation we got, ok.

So, this cannot be ignored actually. This has to be subject to that. So, in other words, we want to find some appropriate selection of P and K e in such a way that it minimizes the steady state P steady trace of P 2 very exact subject to this Ricotta equation basically.

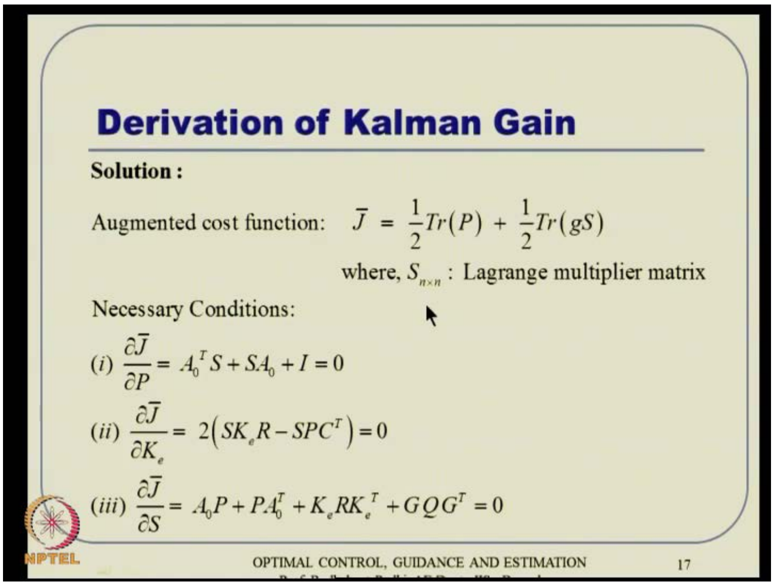
(Refer Slide Time: 33:05)



Now, here is error. Some facts from A from matrix calculus and this matrix, I mean some facts are evolving by it. We actually will not go through the little of this derivation and all. These are all matrices actually and remember, sigma is a matrix and we are all some matrices and somewhat. These result are available will simply that we use it. So, what it tells us that you have got this. This optimization problem here, where you have to minimize this cost function subject to this equation which is equal to 0.

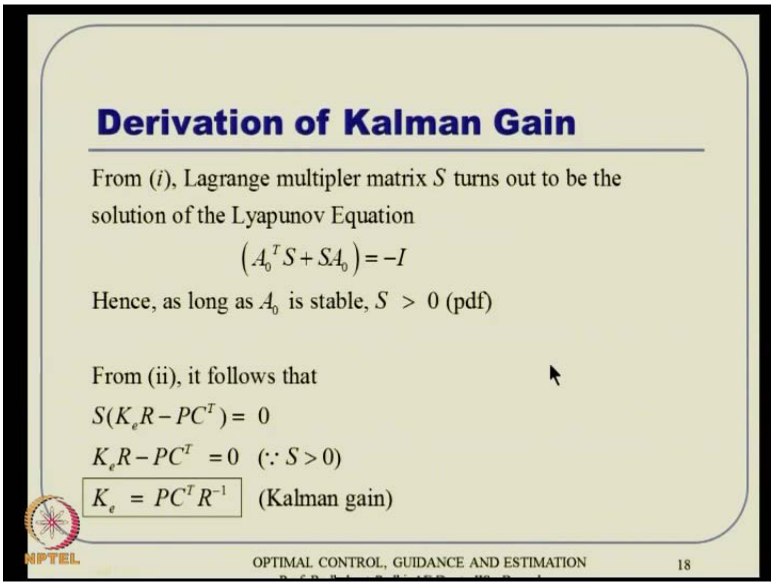
So, the theory tells us that you can have an augmented cost function which is the original cost function plus this half of trace of G times S, where S is a Lagrange multiplier matrix actually. Then, the necessary conditions turn out to that all these derivative. Now, these are function PK e and S.

(Refer Slide Time: 33:30)



So, the partial derivative of J bar with respect to P has to be 0, I with respect to K e has to be 0 and with respect to S has to be 0. Now, this partial derivative can be derived using some of these results here. So, it turns out to be something like this. However, to solve these three equations to gather basically and that is not the typical c. So, from equation 1, you can actually do this. When this equation can be written something like this and this is nothing, but a kind of Lyapunov equation actually, ok.

(Refer Slide Time: 34:22)

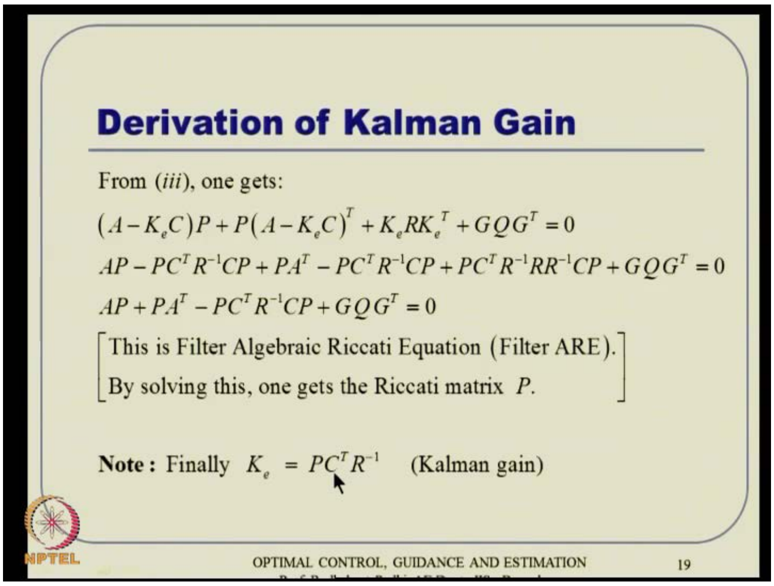


So, hence, as long as A naught is stable, S is guaranteed to positive definite, S naught is a system dynamics matrix for that particular Lyapunov equation and all that. So, as long as A naught is stable, this will be a positive definite matrices which satisfied this equation.

Remember, what is G? It is the equation. S is nothing, but the Lagrange multiply actually Lagrange multiplier matrix and to be like that.

Now, what is you have that let us go to this equation 2. What does the equation 2 tells? It tells something like this. That means K e R minus P c transpose is 0 because S is a positive definite matrix. It cannot be 0 actually, ok. So, K e if you see this equation and this K e is nothing, but P c transpose R inverse. Interestingly, it is exactly similar to what you have actually derived in LQ observer theory. If you can see that of the lecture notes, I mean slides for that lecture, you will see that actually.

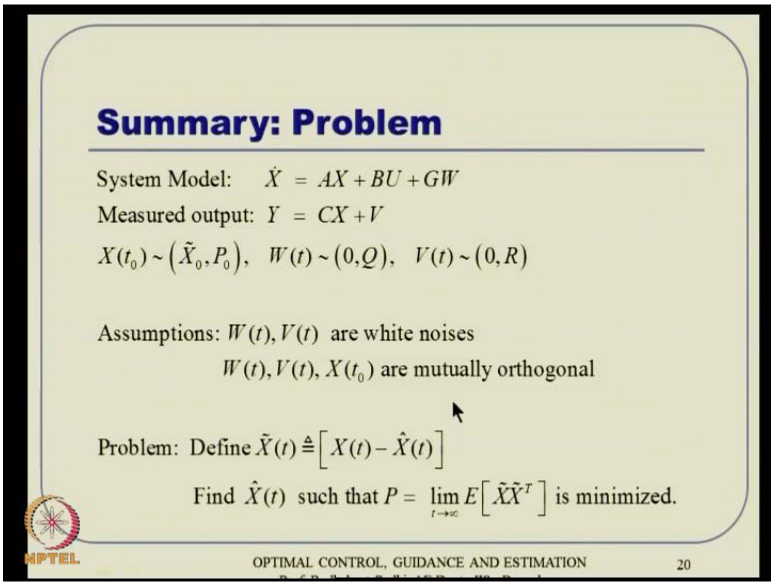
(Refer Slide Time: 36:04)



So, K e is nothing, but P c transpose R inverse. So, we have to use equation 1 to what equation three. Some from three you can put some; you can put it that way. Now, you expand A naught is A minus K e c sort of things. So, that you expand all that is equal to 0 and then, it turns out to be something like that AP minus this P c transpose R inverse c P and all that actually is equal to 0. That means, if I mean you can see that AP is here, then PA transpose is coming from this term K e K e r K e transpose and K e is nothing, but this one P c transpose R inverse.

So, you can substitute that wherever K e appears and get that, then it turns out that two terms will be cancelled out actually. This R inverse is identity. Once this is identity, these two terms are equivalent opposite. So, that two is, they will go. So, you are left out with something like this actually, ok, but this turns out to be the filter algebraic Ricotta equation or instead it is called Filter ARE actually. So, once you solve this, you get the Ricotta matrix P and finally, K e is nothing, but P c transpose R inverse. That is how it is.

(Refer Slide Time: 37:31)

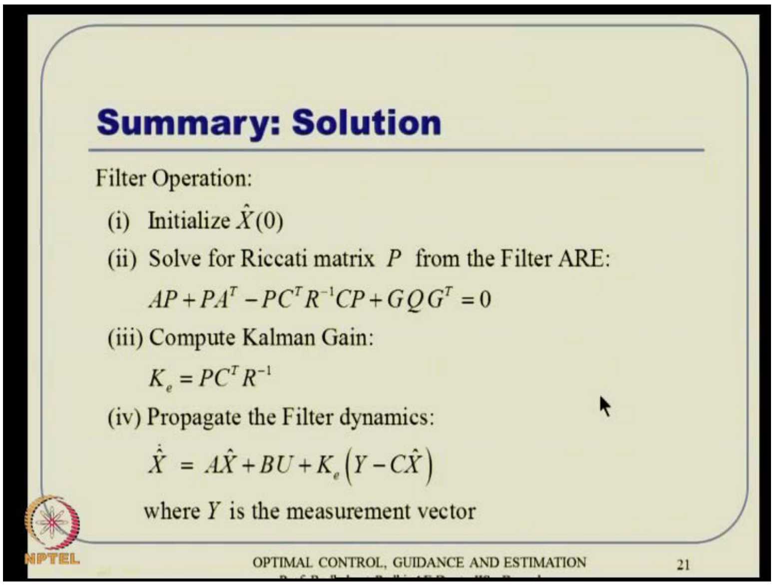


So, summary is something like this. You have got system model which is X dot is AX plus BU plus GW and we have got a measured output which is Y equal to CX plus V and the initial condition you have got process noise, you have got sensor noise. These are characterized by their mean value and covariance matrix. Initial condition mean value is X tilde dot P naught and W is 0, Q and V is 0, R ok.

The assumptions here are these two noises are white noise and the WV and also X naught except t naught are actually mutually orthogonal. You define that error in the state is something X of t minus X hat of t, where X hat is estimate of X and our observer interest is to find the X hat of t such that P in the limit tends. That means steady state P limit t tends to infinity expected value of XX tilde X tilde transpose. That is what it is. So, steady state P has to be minimum actually.

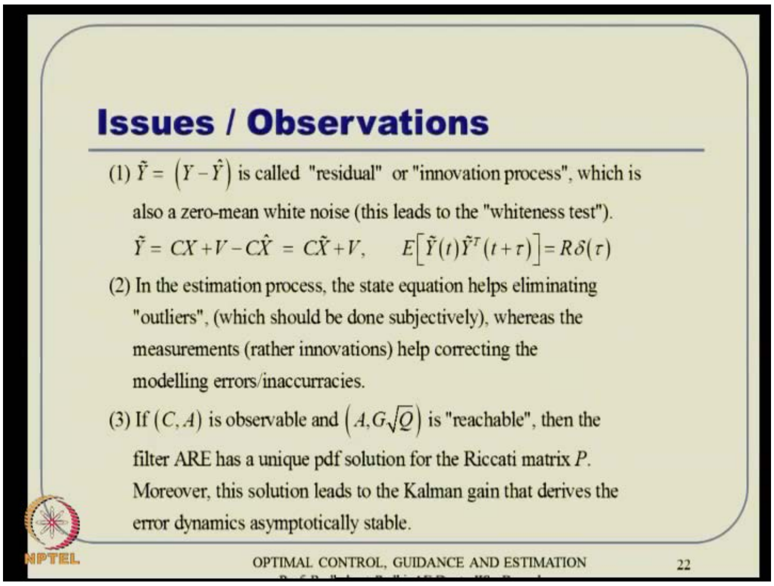
So, filter operation I we initialize it first as something like X hat of 0. Then, we need to solve this Ricotta matrix or Filter ARE that equation which you need to solve. Then, you can compute the Kalman gain and then, you can propagate the filter dynamics. So, the X hat 0 is known and this filter dynamics structure is known with a deterministic value of K e, ok.

(Refer Slide Time: 38:52)



So, they you know nothing unknown now here. So, you can propagate this filter dynamics and because of the nature of the solution and things like that, it is guaranteed to give stable tracking, I mean stable error dynamics. That means error value that X tilde X minus X hat will ultimately go to 0 as t goes to infinity actually, ok. So far so good. It turns out to be very lacerative actually, but there are certain issues or called successful operation of Kalman gain and the certain mathematical comments also basically. So, let us see those one or two actually.

(Refer Slide Time: 39:45)



So, first thing is this Y minus Y hat. How about that is you can think about that is Y tilde and Y minus Y hat is something like well, I should have that yeah Y minus Y hat. So, Y is actual measurement and Y hat is CX hat actually ok. So, estimated output, predictor output whatever you can say actually. So, if you take actual output minus this predictor output, this is called the residual or innovation actually which is also a 0 mean white noise and that actually leads to this whiteness test actually.

In other words, if you implement a filter and then, tell filter is working and you also need to do some sort of few tests to validate your result that it is actually working and we are not misled actually. So, what is the first result tells us that if you think about residual or this one, so it also needs to be some sort of a 0 mean white noise, all right. So, this is what you can do now is you generate this residual for a large number of signals and try to find out the mean value of that and that should go to 0 actually, ok. So, that is called whiteness test and all that.

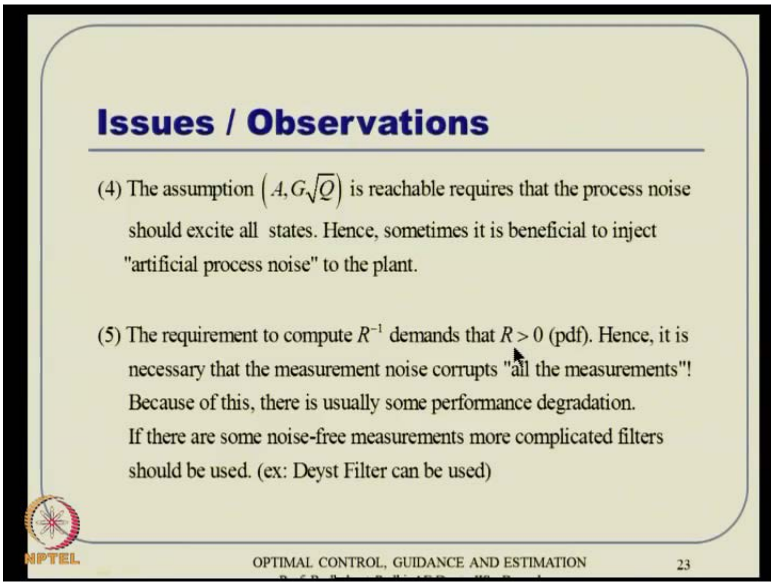
However, this particular Y tilde has to behave like a white noise actually and Y tilde is something like this, CX tilde VCX tilde plus V and then, expected value of YY tilde Y tilde transpose should turn out to be R actually. So, that is actually kind of whiteness test basically. Also remember, in the estimation process, the state equation actually helps in eliminating the out layers. That means, you propagate the state equation or filter dynamics and things like that. You expect some sort of a value for Y hat. You know what Y hat is actually, ok.

Now, if Y turns out to be pretty close to Y hat, then maybe you are wrong or the sensor is good. So, we can operate that. I mean you can update the values and all that for X hat, but if for some reason, some Y happens to be much larger than Y hat, then something is going wrong there. You tell that data particularly turns to be something like a out layer. So, we do not have to use it actually. So, then it tells out that the state equation because you have to propagate to see what is CX hat. So, the idea is here that the state equation actually helps us in eliminating outliers also. That means, if there is some outliers somewhere, then we do not have to operate it actually, ok whereas, the measurements or rather innovations help us in correcting the modeling errors inaccuracy actually. That means it does both the thing. It actually helps us in declaring some third data is outlier and hence, not using it or it actually kind of if you use it because you are relying more on the sensor output and less on the system process modeling actually. So, it actually helps us correcting the modeling errors or modeling inaccuracies as well actually.

So, as the third point, third point is if C, A is observable and this pair A, G square root of Q is reachable, then the filter ARE has a unique pdf solution for the Ricotta matrix P and moreover this solution leads to Kalman gain that derives the error dynamics asymptotically stable. So, the two conditions. One is it has the problem to be observable, otherwise nothing will work. Your Kalman filter is also not good work actually. So, for the Kalman filter to work, this pair C, A has to be observable first and this pair A, G times square root of Q needs to be reachable actually.

If this condition observe, its observality is not that. You can do anything. You have done top of observality. This has to be there for unique solution of this filter. If it is there, then we can get something like a unique p d f solution or positive definite solution from algebraic Ricotta equation or filter ARE equation, ok and moreover, this solution will ultimately lead to Kalman gain. Remember, K e is nothing, but Pc transpose R inverse. So, ultimately it will lead to Kalman gain that derives the error dynamics asymptotically to 0 basically, ok.

(Refer Slide Time: 45:20)



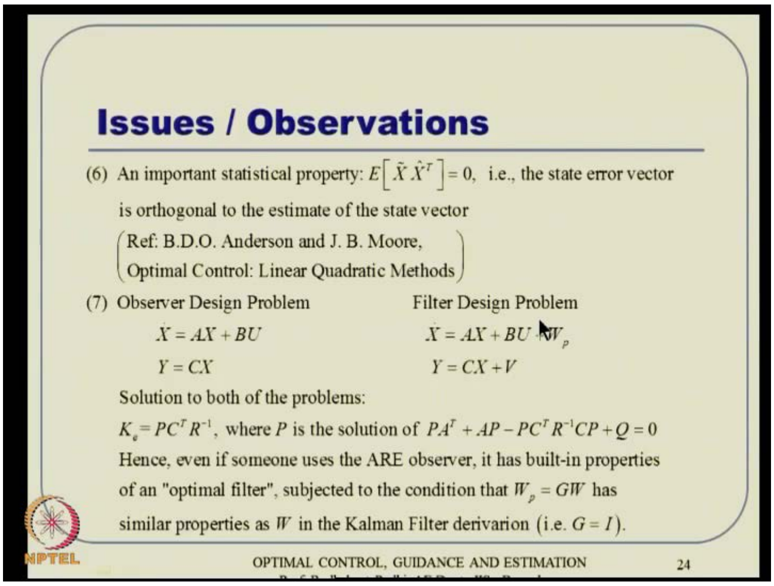
Now, the fourth observation. The assumption this one, A, G times square root of Q is reachable requires that the process noise should excite all states actually. Process noise keeps on coming and ultimately should excite all states and if we does not tell properly something, somewhere something goes on and you will not be able, you will never be able to observe it properly actually, ok. So, because there is a requirement that you should excite all states sometimes, there is a necessity now to inject artificial process noise to the plant. For example, if you talk about kinematic label and dynamic label variables, typically system dynamics contains two label variables, something like kinematics which is something like X dot is V and dynamics, where V, where V dot is not A, but A plus W. We can say that way, ok.

So, because this is a physical quantity normal escalation, right, I mean escalation applied to the system basically. That is directly influenced by noise W, ok. Now, the question is because it needs to excite all states, now W is suppose it does not excite explicit, give some exciting V dot V only through this V dot equation, then it makes to a kind of inject artificial noise on that. Remember, X dot is V simply by definition is no one wise and nothing that actually, but still by putting an artificial noise, the W1 things, the filter can work beautifully actually, ok.

So, that is what it is. Then, another starts. The other point is the requirement to complete R inverse demands that R is first definite and hence, it is necessity that B, the measurements that all we put all the measurements. So, with that measurements, that R inverse has to happen because we cannot define some sort of measurement till that particular component I do not know what is going on actually, I will tell that. So, what it happens is the requirement to complete R inverse demands that R needs to be positive definite N. Hence, it is necessary that the measurement of noise corrupts all measurements and because of this, there is usually some preference. I mean there is some performance degradation and if there are some noise free measurements typically either that way. That means, you artificially inject the noise and make everything obey there. So, they will leave it like that or there are some noise free measurements or some little bit more practical complicated filters should be used. They are typically there in the literature. One thing comes from the deist filter. Deist filter for example, actually.

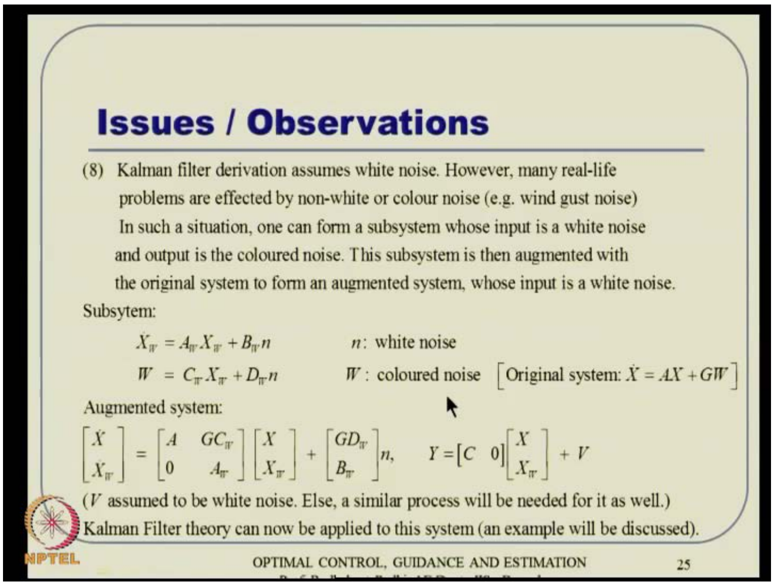
So, there are bunch of issues here for successful operation of Kalman filter. We will see more on that when we finish external Kalman filter and later actually, all right. So, only then, if you make sure that these things are not forget and these are kept in mind, then you can get nice behavior of Kalman filter actually. The other one which is another statistical important statistical property that expected value of X tilde X tilde transpose should be 0. That means, the state vector state and error vector, this is this estimated state is estimated error.

(Refer Slide Time: 48:44)



So, the state vector is orthogonal to the estimate of the state vector actually. We can see, understand more common details and lastly, this comparison sort of things. Suppose, we have observer design and filter design, that means, you have this X dot is AX plus B here, but there is a process noise plus W and you have got Y is CX and you got Y of CX plus V here and testing the solution to both problems turns out to design that is, I mean K e equal to P c transpose R inverse and where P is the solution of this Ricotta equation actually.

(Refer Slide Time: 49:58)



Here, if you talk about other one, it is Ricotta equation contents G also GG transverse. This side also actually from then this part I have already told before Kalman filter derivation assumes white noise and some many real life problems are typically not detected by white noise. Rather color noise are very frequent, where say for example at all wind gust noise and all that actually.

So, in that situation, one can form a sub-system whose input is a white noise and output is a colored noise and explain all that little bit before and the sub-system is then augmented with the original system to be form. An augmented system whose input is white noise actually. So, this is what it is. Initially, this I showed an example already before actually. Usually what happens is, you have got a transfer function or you have got some sort of a small time state space model for which you take input as white noise and output as the actual colored noise that you are expecting actually. So, then what happen is you tend to kind of augment these two together and then, you will get whatever you want actually.

So, this example I have just given little while before actually or may be in the previous lecture you can see basically. So, this is a need concept because many times, it is actually colored noise, but colored noise can be actually a dynamic model or static model, either way actually. You can think about putting that some sort of AAA system dynamic associated with that colored noise is something like this. So, you have got AXW dot is something like this, RN is a white noise where W is a colored noise and that is a function XW and white noise both.

So, you can put that in the state space form and then, carry out all that basically. Now, if you do that, the Kalman filter theory now can be applied to the entire augmented system basically. So, then it can be addressed and this is typically called stuffy filter ideas and all that each other actually. If you see that there are bunch of these comments are here issues, observations, comments and all just be aware of that. There are many more comments and many suggestions in the iteration as well actually and will slowly go that is discrete time derivation. Then, how do you put continuous time discrete times together. That means, system dynamic can be continuous, but measurements are discrete.

So, how do you do that? Finally, we will extend the same concept in the same discrete, time discrete continuous setting for an external Kalman filter which is very heavily used and with lot of applications. Then, there also we will talk some of these issues and recommendations, something like that actually, all right. That is what I wanted to talk here. Thank you. Bye.