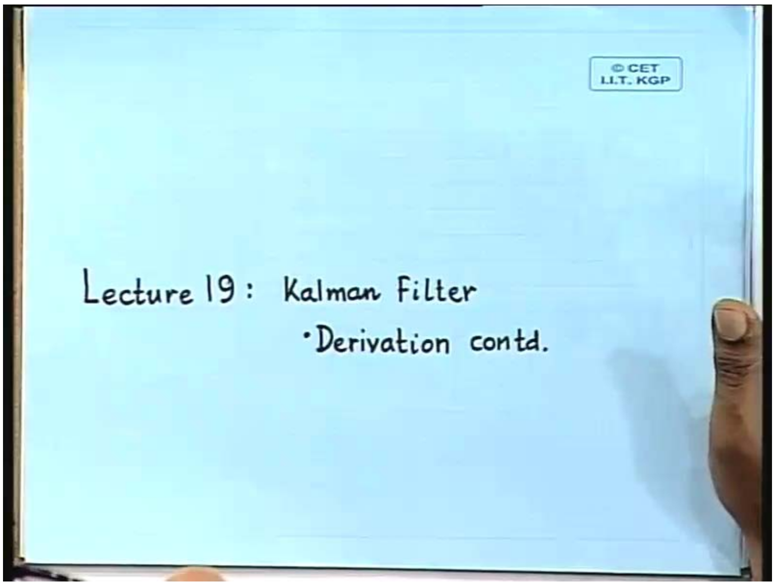
Estimation of Signals and Systems
Prof. S. Mukhopadhyay
Department of Electrical Engineering
Indian Institute of Technology, Kharagpur

Lecture - 19
Kalman Filter – Derivation Contd.

From where we left yesterday, so what did we find yesterday? Yesterday we had seen, basically two steps to remember that; we first solved the least square problem and found a solution, then we which is the which was the weighted least square problem, in which we choose some weight matrix W. Then we found then then from among that solutions we we found that, if we choose W is equal to some R k inverse, then we we find that the variance of the states is minimized, right.

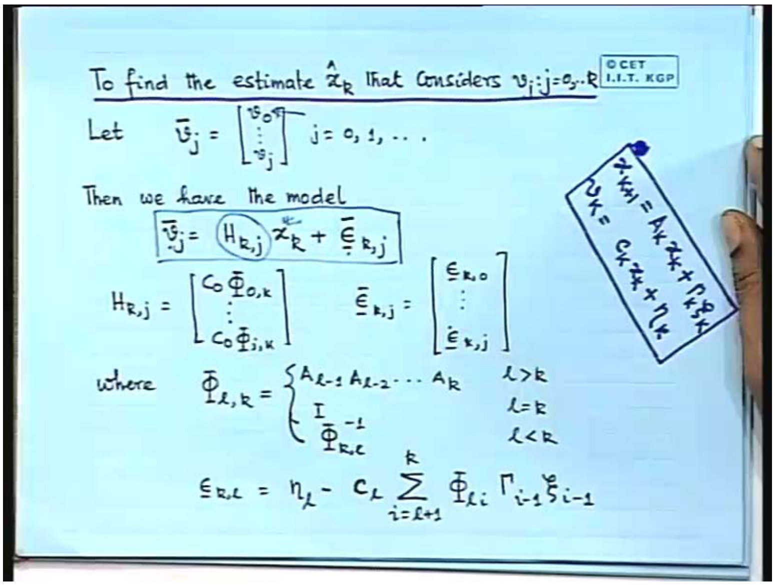
(Refer Slide Time: 01:33)



We did this for a single measurement v, now remember that; even if it is a single measurement, firstly remember that it is a vector and in our treatment, nowhere we it is for a general vector v. It is so happen that, it it it was v k but we can do it with respect to any other vector. The analysis are as such did not assume anything, it just found simple least square solutions, may yes to a linear vector optimization problem, nothing else.

So the idea is that, if we want to do the same thing with respect to all the measurements from v zero to v k, then we can all we have to do is that, we have to formulate a similar estimation problem involving measurements from v zero to v k. Exactly a similar linear vector measurement problem, and then solve it in the exactly the same manner. So that is what what what we are going to do today, okay. So first thing, that we are doing, now previously we had found out x hat k, that consider only v k in the last days notes class.

(Refer Slide Time: 02:54)



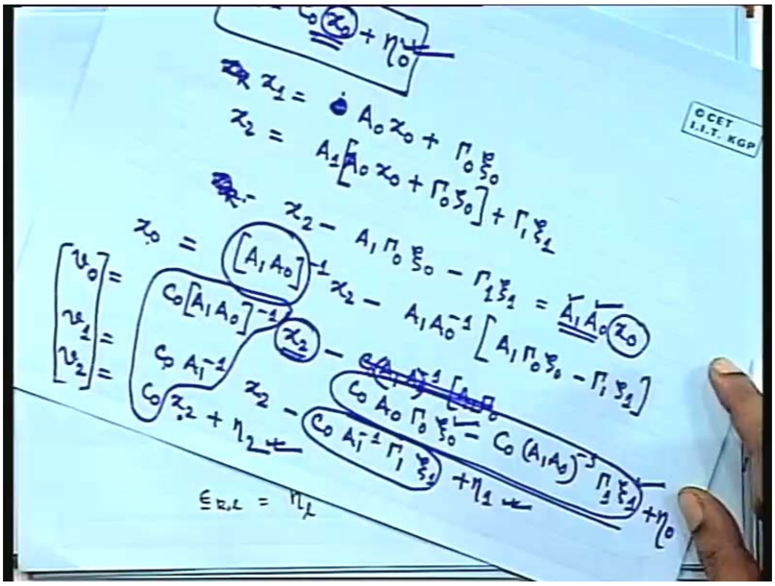
Now we have we have we will find an estimate, which will not consider only v k but we will consider all v k, v k from v zero to v k, right. So that is why we will solve a least square problem, which will minimize the weighted sum of all these errors and then be be be be minimum variance. So now first of all we have to write the system equations in a, previously what was our equation? Our equation was v k is equal to C k, x k plus xi k. So v k so the left hand side the is is the measurement vector, then a matrix C k multiplied by the thing that we want to optimize x k; you want to optimize an estimate of x k plus some noise term, this was our form.

So first of all when we consider so many measurements, you have to again pose the problem in that form, first. So so now how do you, so it it turns out that; if you write this v, you remember that in initially in the last class, we did it with respect to a single element. Now we are cascading all the element, these may be individually these may be vectors. So we are putting one vector, then the next vector, then the next vector; we are making a long vector, okay. Now how do we now first of all we have to write our, to be able to utilize our last days results; we have to cast the problem in such a form, where we will have the measurement on the left hand side as a vector, then we will have some matrix, multiplied by the quantity, that we want to estimate plus some some a vector, you have to first put it in this form.

So so how do you put it in this form? So we apply the normal state transitions rules; that is what we are trying to say is that, we can always write that let us say, we can always write like this. That is what is the state transition relationship? Say our equation is, what is our equation? Our equation is this v is the, in this case this v is the, one second our equation is like this, x k plus one is equal to A k, x k plus gamma k, xi, k, okay. So if it is like that, then then you can you now now and finally; we have v k is equal to Ck, xk plus eta k, this is our system model, okay.

So now what we have to write is that, we have to cast we have have to express these in terms of in this form. How do you do that? This is again a kind of algebra, where so what we are what we are writing is, this is you know this is a kind of symbolic form. This what does it say? It says that, I have difficulty in writing, it says that v zero for example, let let's take the first term; first term is v0, right.

(Refer Slide Time: 07:23)



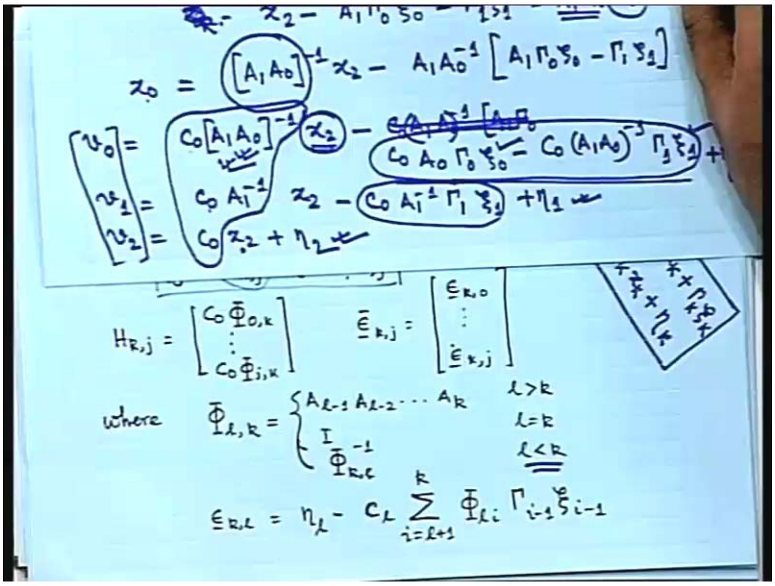
v0 is equal to what? v0 is equal to c0, x0 plus eta0, okay. Now if we have if if we have a concept of a state transition matrix, then we have what do we have? That is how do we write, x k in terms of how do we write x k in terms of x zero? For example, we write x1 is equal to x one is equal to phi is equal to A0 x0 plus gamma0 xi0. Then we write x2 is equal to A1 A0 x0, x two is equal to A1, x1 plus gamma1, xi1, right. So in in other words, if you if you go on doing it like this you will get that, you will get x2, xk is equal to what? One second, you will get, right, correct. So you will get x2 minus A1, let us get x2 from x0; just let us put k is equal to two to understand. So we are getting x2 minus A1, gamma0, xi0 minus gamma1, xi1 is equal to A1 A0 x0, correct.

Now how do we get x zero? What what do we have to get? We have to we will get this one from x k, remember that we are estimating x, x at at time instant k, but we have to relate it to the measurement at zero; we are going we have to go back, okay. So so so now here is an assumption, that all these matrices are invertible; this we are assuming again as I said right in the beginning, that we have to assume certain things. So what we will this be? This will be A1, A0 inverse x2 minus A1, A0 inverse, A1 gamma0, xi0 minus gamma1 xi1, correct. So now what is A1, A0? Now you can say that, A1 A0 is a now you see that; so you see that, x0 is equal to this. So what is v0? So v0, I put this x0 here. So I will get C0, A1, A0 inverse x2 minus A1 C0 A1 A0 inverse A1 gamma0 xi0. Now A1, A0 inverse means; A0 inverse, A1 inverse A1 inverse into A1 will get cancel.

So actually this term will be C0, A0 gamma0 xi0 minus C0; this will be A1, A0 inverse gamma1 xi1, correct. This is so now you see that, if you if your k is 2, then you have expressed v0 in terms of x2 and the noises xi0 and xi1, correct. Similarly, you can write v1, similarly you can write v2. So if you now now you can form a vector v0, v1, v2 and everywhere, you are you are going back to x2. So in this case what will will be? It will be C0, A1 inverse x2 minus, C0, A1 inverse gamma1, xi1. This will be what? This will be C0, x2, there is v0 is plus eta0 is there, if you write v0, you have to write plus eta0. So plus eta1 plus eta2, because of this term. So you see that you can it it is now possible; now now now you can now, this becomes your h matrix. So it is h h into x2, plus some term which is a linear sum of the process noise components and the measurement noise, right.

So these terms these terms, I have made an this is the basic idea. So now therefore, these terms can be expressed in this form, you see what is this H k j matrix?

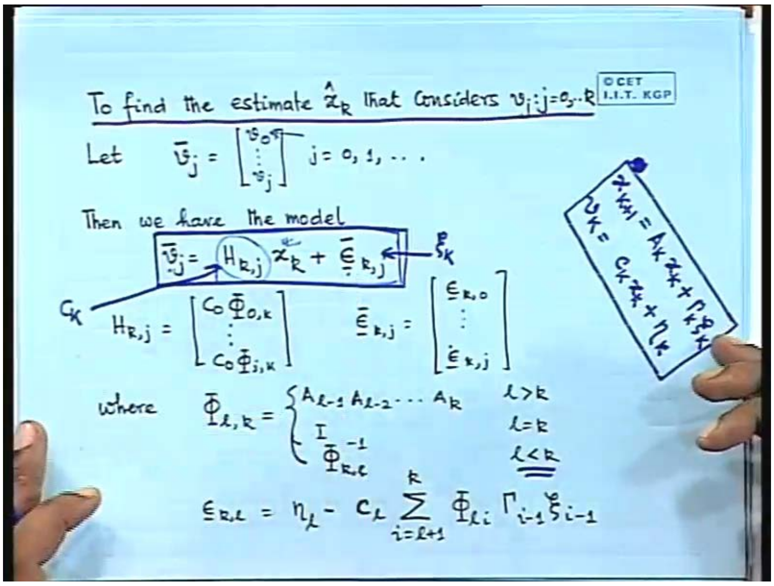
(Refer Slide Time: 13:27)



It is c0, phi0, k. So what is phi0, k? phi l k has been defined as from A k to A l minus1 where l is greater than k. Now in this case A is less than k, because it is 0k, k must be greater than zero. So phi k l, so phi l k when l is less than k is equal to phi k l inverse. So it is nothing but phi k0 inverse. Now phi k0 is what? phi k0 is A0, A1, A2, Al, Ak whether Ak minus one inverse. So you see when when when you have k is equal to 2, then you have A0, A1 inverse, right.

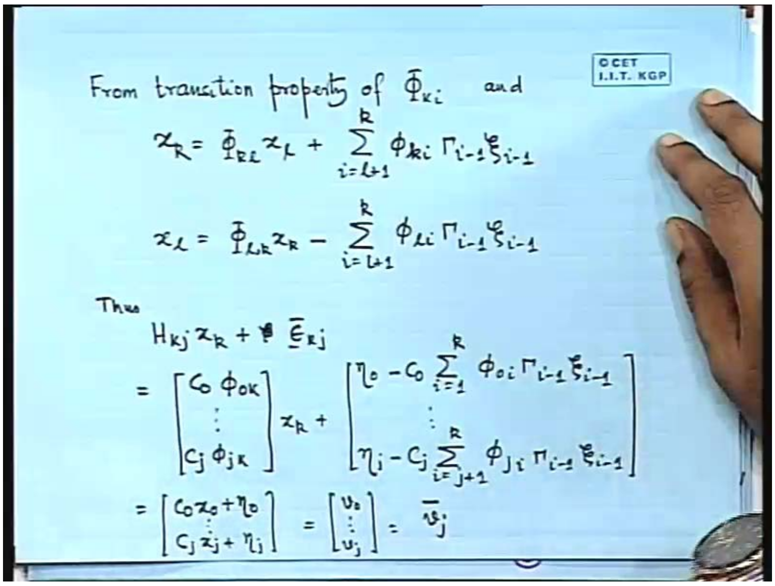
So in this way it will go, next one will be one inverse, next one will be i. So these are you know, kind of inverse state transition matrices, if you know what are state transition matrices. This is this is just algebra. I mean mean I I mean I just wanted to show that, you it is it is it is indeed possible to be cast in that form. So the so the major thing is that, you can cast you can write this problem into this linear form; in which we have already solved this problem.

(Refer Slide Time: 14:45)



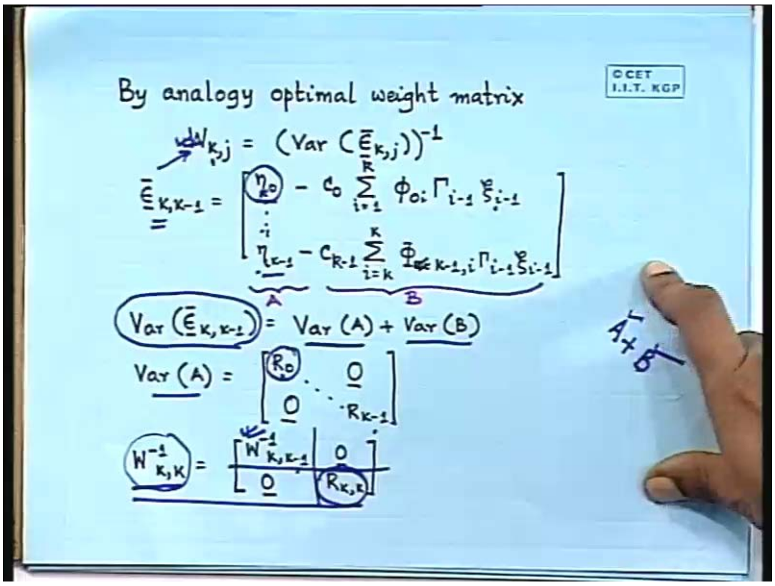
But this will now if you if you now solve this least square problem, and then choose our weight according to the according to the; see previously what what happened? Here what did we have? Previous problem, we had here we had xi k, if you recall and this one was our Ck, correct. So now so now we have we we are trying to solve this this augmented problem. So so we will do exactly the same thing. That is after all the least square solution, I mean does not depend on the interpretation as long as you have a problem in this form; you will get the same least square solution, is it not? So therefore now we can use our last result, that is this is just what I proved, just right now, the same thing written in written in general notation.

(Refer slide Time: 15:40)



So x k is equal to this, I did exactly that for for that for that one, two, case. So this we can skip, probably now.

(Refer Slide Time: 16:01)



So now now we know, now previously what did we see? We see we saw that, if we choose Wk j equal to Rk j inverse, we get the optimal solution. What was Rk j? Rk j was variance of these, in our last this problem. So now we have to choose variance of this as the optimal weight matrix. So we have to choose Wk j as variance of this term, there is a lot of algebra here today; we cannot help, it it is probably the simplest case, we cannot so so so please bear with me and and try to understand. Now this term is is of this form, as we have already seen that it will contain all these eta, one eta term and the rest will be all a linear combination of this past xi terms, is what it will be.

So so so there are two terms, now the question is that; if this is epsilon, what is variance of epsilon? So we want to find out this W, imagine that what are we trying to say? We are trying to say, that is a stochastic vector A and if there is another random vector B, what is the variance of A plus B? That is what we are trying to find out. So now can you imagine, what we what we are trying to do? You have a vector and we are trying to, so you know one one interesting thing about matrices; is that if you can if you if you break them into sub vectors, and if you can write it is as a block matrix. For example this itself is a vector, but just for the time being treat it as an element. It is a vector but just treat it as a scalar element. So then this matrix has this element, this element this element, as if they are scalars then if you just multiply, then when the same same kind of results will follow.

So you do not really need to treat them as vector and and get more complicated. So so if you have, what is variance of A plus B? In this case it will be variance of A plus variance of B, why? Because these two vectors are independent, because here we have.. what were our assumptions? Our assumptions was that eta and xi are totally uncorrelated. So so any of these terms, that is this and this no correlation and so now what is what is variance of A again? We have assumed that, each one of them are actually white. So therefore eta0 is not correlated with eta1, is not correlated with eta2. So therefore when you take variance of A, you get individual variance of eta0, individual variance of eta1, their cross variations are zero, even eta zero, eta one does not have any cross correlation. This is our assumption in fact to make them in this form; we have made those assumptions, otherwise I mean the algebra becomes too complicated to solve.

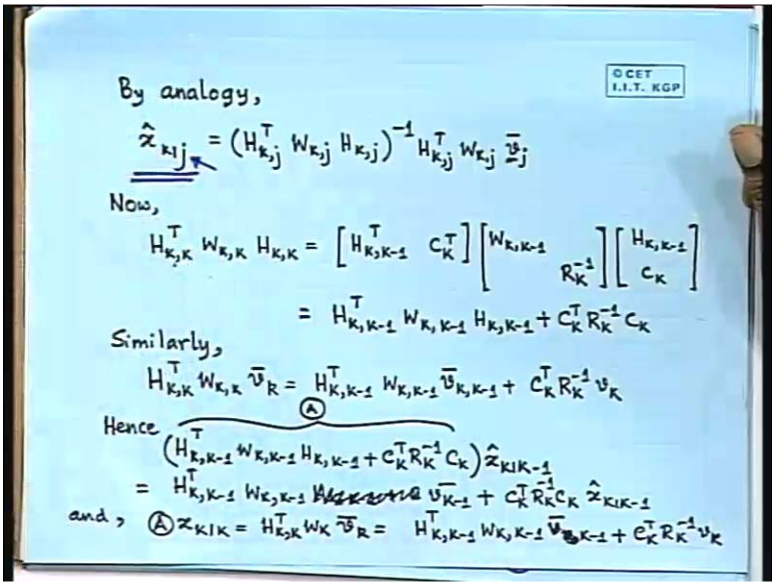
So variance of A comes in this block diagonal form, these are these are individually matrices; because these are individually vectors, you have to visualise this, okay. Now Kalman filter is actually a recursive algorithm. So we have see how we we are not solving, you know what is the problem of solving this problem? The problem of solving this problem is that, as k goes increasing; this vector becomes longer and longer in length. So after sometime it will become so long that, you will not be able to compute the solution. So we are not going to compute the solution that way. So we have to have a finite fixed dimensional problem; whatever equations we have must have constant dimension vector, we cannot our problem in which the vector dimension is increasing with time. So therefore we have to recursify it, that is we have to compute our new solution; based on our earlier solution and some constant computation. Unless we bring any algorithm in that form, it is not useful to compute in, I mean especially in real time. So therefore our objective is that, we have to we have to recursify this form; we know the so we know the non-recursive solution, but it is not useful right. So we had that is what we are trying now.

Now let us see what happens? What is this remember what is this Wk j? Wk j is the weight matrix, when you try to solve for xk; taking care of measurements from v zero to vj, that is why we called it Wk j, correct. Because on the left hand side we have v zero to vj and we are trying to solve for xk. Now let us see, now we are trying to we we will go about ((recursifying)) ((00:21:31 min)) it, so we will first note that first note this identity. This is pretty cool, because after all WA is this inverse; so therefore W inverse is nothing but the variance. So what is W inverse k k? W inverse k k means, I am trying to solve for xk taking measurements from v zero to vk same instant; and at every instant my measurement vector is increasing by one, correct, correct. Some v zero to vk minus one, then I am adding one vk term at the bottom it becomes a longer problem, okay.

So if I take the inverse of this it will become a longer matrix, I mean the variance of this; because with the error term also one more one more block rows will come. So when you take the variance of that; then you will get for the first n minus rows, for the first n minus one rows you will get the k k minus one problem. So k k minus one, k minus one means you are trying to take care of measurement from v zero to vk minus one. So you have v zero, v one, v two, vk minus one. Now on the left hand side you have added another row vk. So when you take when you multiplied by its transpose, then you will get the block matrix and you will get one matrix corresponding to vk, vk transpose. That is the rather rather eta k, because a new row eta k will now appear here. So corresponding to that, you will have Rk, k.

So first thing is that this can be written like this, we have a long way to go out so do not shy okay. I have tried to make it simpler than the original source, from where is which it is taken now. Remember our old solution, what was our solution? What it what is our solution?

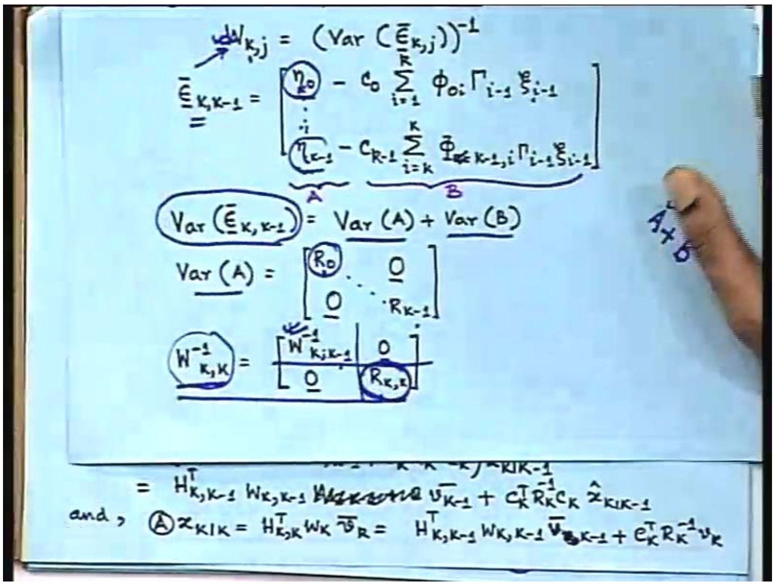
(Refer Slide Time: 23:35)



Our our solution, if you recall was finally after putting that R k inverse matrix. Our first problem taking only vk was C k transpose, R k inverse, C k whole inverse, C k transpose, R k into V k that was the solution. So the so I have just put the corresponding ones, this is corresponding to C k. This is corresponding to R k inverse, W k. So I have just put those. This is a this is C k transpose, R k inverse, C k whole inverse, C k transpose, R k, R k inverse v, v k; this was our solution. Just put them in the because, it because this is same problem longer dimension.

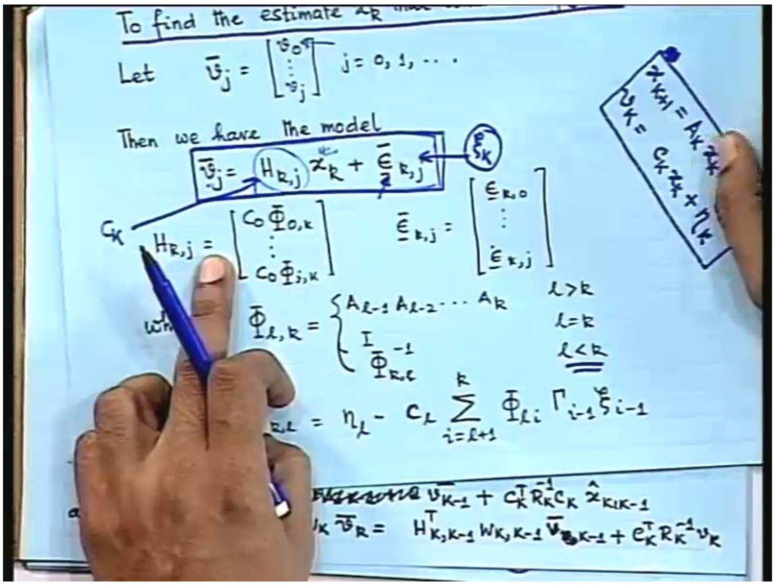
So this is my x hat k j, that is the optimal estimate of x k; taking into factor measurements from v zero to v j, this is my general solution. Now I have to recursify it, which means that I have to go from k minus one to k. Then again from k to k plus one then like this I have to go, right. So essentially I have to put as j, I have to put k minus one and k and see how are how they relate? So that is very simple. Now from from the previous from the previous two formulae that we have already obtained,

(Refer Slide Time: 24:52)



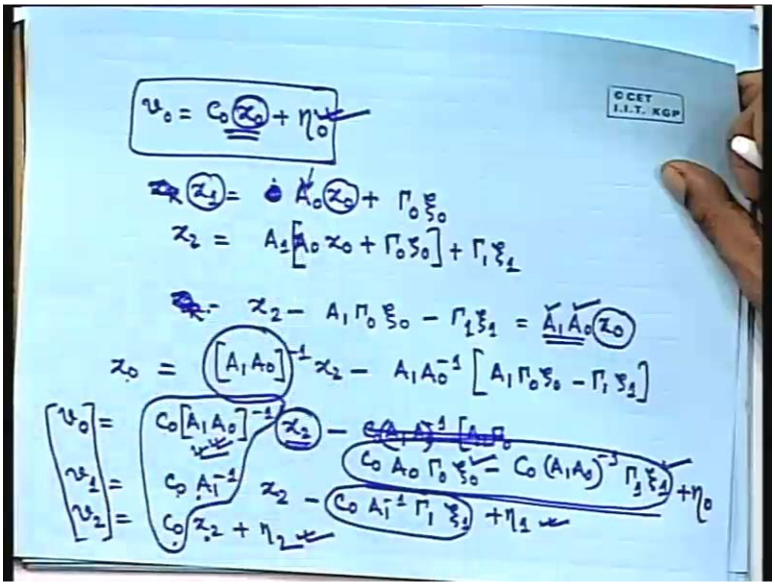
that is this one equal to this and H k, k is nothing but H k look at H k, k

(Refer Slide Time: 25:00)



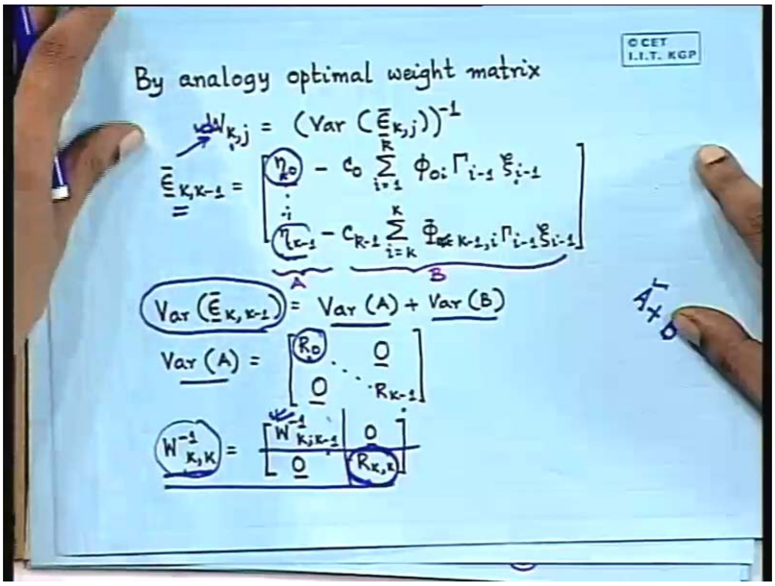
what is H k, k? H k, k will be C0 phi0 k, C0 phi j k. So if you have H k, k you will have how many zero k? one k, two k, three k, up to j k. So in in this case, it will be k k, if you have k k k minus one; you will have one row less, correct. So so therefore; you can always partition the matrix like this and and what is phi k, k? Rather phi k, k is one because from from to to take the state estimate from k minus one to k; you need to multiply it by the matrix A, but is it not what were we doing. We were, see, we have we were multiplying

(Refer Slide Time: 25:52)



See if you want, if you want to take x0 to x1, you have to multiply by A0. If we have to take xk to xk you have to multiply by the identity matrix. So phi k, k it it is the it is the state transition matrix, I think so between k and k the matrix is i. So therefore, so therefore here you have only C k, no phi because phi k, k is one. So you can partition the matrix like this; H k k is this one plus one addition row, in this case one additional column because this transposed actually it is like this C an additional row. And Wk, k we have we we have already partitioned, that will come like this; because if if this is W inverse k, k then then Wk, k will be just again just treat them like a diagonal matrix as if these are scalar elements.

(Refer Slide Time: 26:40)



So it will be the inverse of this and the inverse of thi, if you the these are the these the I mean usefulness of block matrices. So then Wk, k is this this this inverse. Now you multiply this and you get, simply just multiply a diagonal matrix. So you get this term, so this term see see why why I have tried to relate them? What is if that I have got, if I put x hat k, k I will get all these terms. That is if I put j is equal to k, I will get this term here. Now I have to relate it with with j equal to k minus one. So if I put j equal to k minus one, I will get this terms here.

So I have to I must be able to relate them, that is why I am trying to relate them; that that how this matrix relates to this this solution. So I have related and then found that, that this is the relationship and this term appears additional. I must go from k minus one to k gradually; as I get new measurements as I want to go in advance, that is from xk to xk minus one. So all those I have to do, okay. And so this part I have taken care of, now I have to take take care of this part. So this part now I am writing; exactly in the similar way that this part will be like this, again simply by partitioning multiply only this to this. So you get this into in to plus this into this, that into vk minus 1. And and total thing is in place of this vector, put this vector. This vector is VK minus1 on top and small vk at the bottom; so multiply these three you will get this, so now this part I have related here and this part I have related here, okay.

Systematically, only this this that they are a few matrices; we have to carefully keeping indexes, maintain, you have to go and multiply. So now if you find, now now you put it put it here, that is this whole… So now how do you get this? If you put k, k minus1, put j is equal to k minus one, j is equal to k minus one, j is equal to k minus one and then on the left hand side; if you.. say this is assumption number two, what was assumption number one? That all the A's are invertible, that was assumption number one. Now here is assumption number two that this is invertible. Previously also in in our last problem also; we we assume that that Ck transpose, Rk inverse, Ck is invertible same assumption we made. See if this invertible, you can pre-multiply this by this. So then this thing, you have this thing you are pre-multiplying with this and you are adding this term. Now this into this from this equation; so pre-multiply this for j is equal to k minus 1. This will come here j is equal to k minus1 will be equal to this.

So this into this is this term, and and and this is simply added, correct. Now this whole thing is, what? This one, so this A this total term I have called A; so this term is A and wait wait wait, correct. Now this is if you put j is equal to k, then this term becomes what? k k k k k k everything will become k, all the j's will become k. If you want to solve for x hat k, k if you want to solve for x hat k, k minus 1 in place of j; everywhere you have to put k minus, this is general equation. So similarly you can write that, just like you have written this one; that is this inverse, this is equal to this. So you can write that same equation, either for j equal to k minus one or for j equal to k. If you write it for j is equal to k, then you wil get this inverse, this into this is equal to this for for j is equal to k.

Now now this, so exactly that is what I have written. Now this equal to A; see this one from this equation, this is A the same A is here. This whole term is A, so what you are doing is so you have and so A is equal to this. So A into xk, k is obviously this one; that is same equation I have put j is equal to k, this is the same equation I have put j equal to k. And that is equal to that is equal to what? That is equal to, this one I have broken up, no no no I am getting, I I

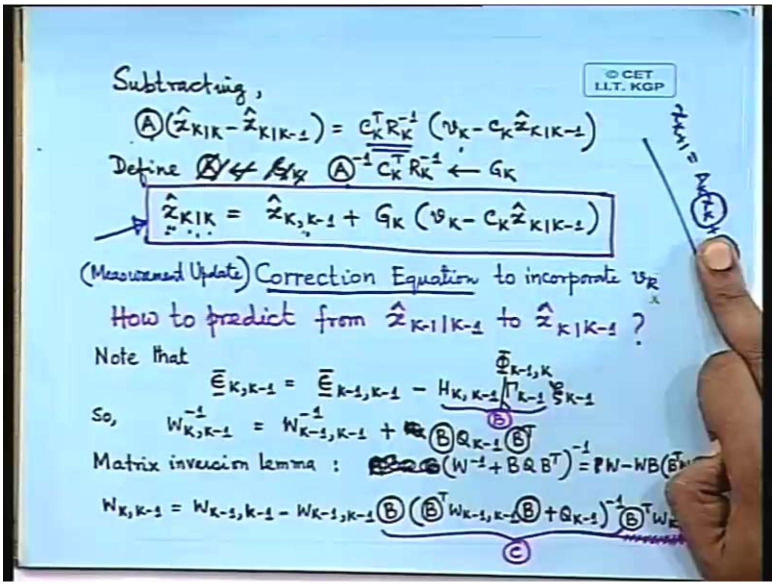
Student <first term on the right hand side k, k value>

[Conversation between Student and Professor – Not audible ((00:32:50 min))]

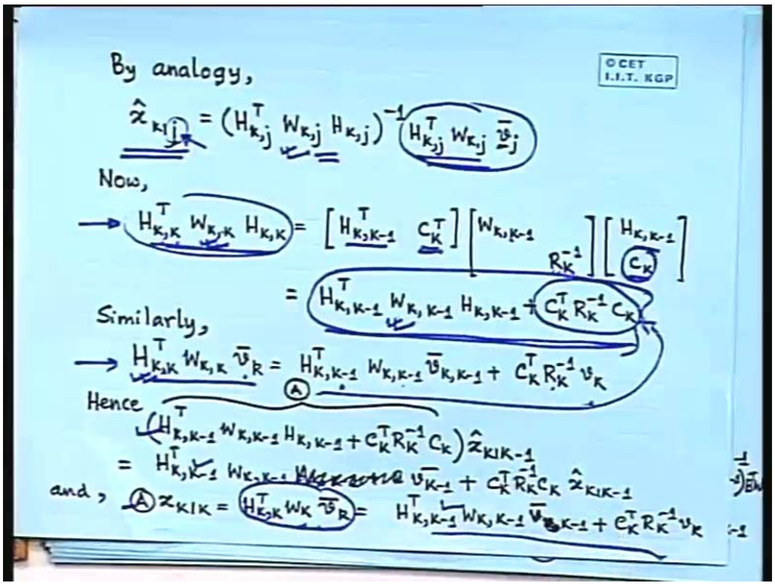
This one, this one, it is this one. See the I got this one and this one I have already said that, that this one equal to this one. So I have just put it here, now I am going to subtract. Here I have got A into xk k, given k minus 1, here I have got A into xk given k; have to finally express xk given k in terms of x, k given k minus one, what am I getting? This is called the measurement update. That is I am estimating xk based on measurement from v zero, v one, v two, vk minus one. That is my estimate h hat k given k minus one, given means given measurements up to. Now I get another measurement which is vk; so I have some more information, so ideally speaking I should be able to estimate xk even better. So how do I now update my previous estimate x hat k, given k minus one, after I get a new measurement and make it x hat k given k, this is what I am trying to do.

So I am trying to generate x hat k given k from x hat k given k minus one, after another set of measurement of Vk becomes available to me. That is what I am trying to get at, okay; because as I get more and more measurements, I should be able to refine my estimate more and more. So I am that, that is why so now so exactly, so now you subtract this and this; that is subtract this from this simply subtract, which terms will get cancelled? These two term will get cancelled, so you will have this A that is that old term into this is equal to so so so these two terms; as I said will get cancelled, this one will cancelled with this one.

(Refer Slide Time: 34:57)



(Refer Slide Time: 35:04)



So they are same only these two will remain. So you get this and then then from this you can just invert the other one; so you can write x hat k given k, is x hat k given k minus one plus all these. That is this inverse this, you have to we are basically multiplying by this inverse on on on both side pre-multiplying; so then this will go off. This will this inverse, A inverse will come here this one, and then take x hat k given k minus one on the other side; this is the first equation on the Kalman filter, which says that if you had a measurement, if you had an estimate of xk given measurements up to k minus1 and if you get a new measurement vk, now. Then this is the way that, you must update your your your your estimate; to make it now optimal with respect to the with respect to the set of measurements, from v zero to vk.

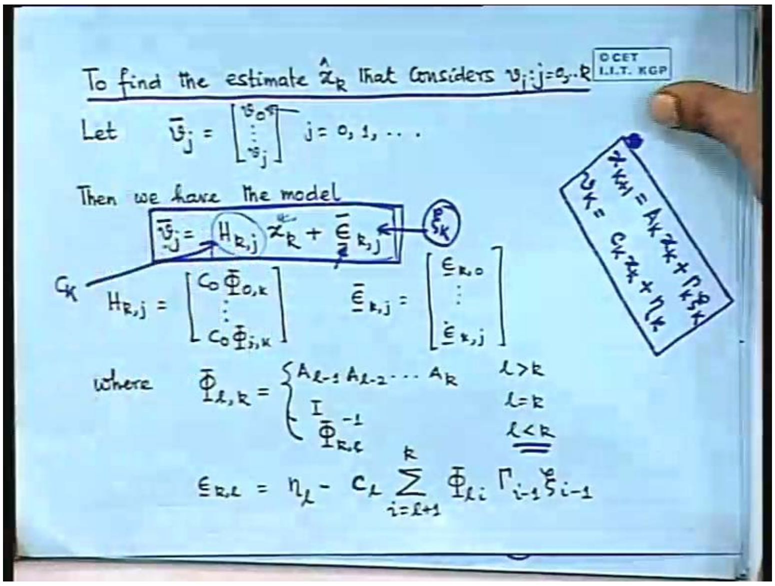
So basically what I have done? I have solved one block problem, I have shown that if you solve this this least square problem based on this block thing; first you solve it for k minus one then you include the new measurement, it becomes a longer vectors, longer matrices then you solve it for k. If you get this two solutions of the same thing xk, then these solutions are actually related like this. This is what I have proved by lot of jugglery with matrices, nothing else. It is just algebra plains taking, tedious, but simple algebra there is no big concept in it.

This is sometimes called the correction equation, because actually the Kalman filter is supposed to be divided into two parts; one kind one part is called the prediction equation where you go go from xk to xk plus one that is you that the estimate, at at which instant is the estimate that you advance. Another is you you relate two estimates at the same time step but which take measurements, which which taking a new measurement into account. So one is a measurement update which is called a correction equation; in the in the sense that you first obtain an estimate of xk, without any measurement at the instant k. So only up to k minus one then once you got that, then you you you you correct your estimate, once you get a measurement at k. So that is why it it sometimes called the correction equation and sometimes called a measurement update, because this equation is executed once you get the measurement at vk. So it is a measurement update.

So I have related this to this, but remember that we have still; that is my estimate is at the same instant k, but if I have to go on doing it then I have to advance k. I have to first make a make a make an estimate of x one, then I make to then then I make an estimate of x two, then make an estimate x three, so have to advance the first index. So how do I advance the first index? I have to advance time, okay. Here I did not advance time, I am.. left hand, right hand side I am I am standing at k. So advancing time is is again simple; involves matrices there is one see interestingly, I mean there are there are there are two approaches one one thing is going again rigorously, see we have I mean one one good thing is that we have solved the problem for for arbitrary k and j, on the two sides of this this problem this this this basic problem that we have formulated.

Fortunately k and j are actually unrelated in the sense that, j can be anything between zero to k. So so that way, I can I can write this problem for some value of k and some value of j.

(Refer Slide Time: 39:33)



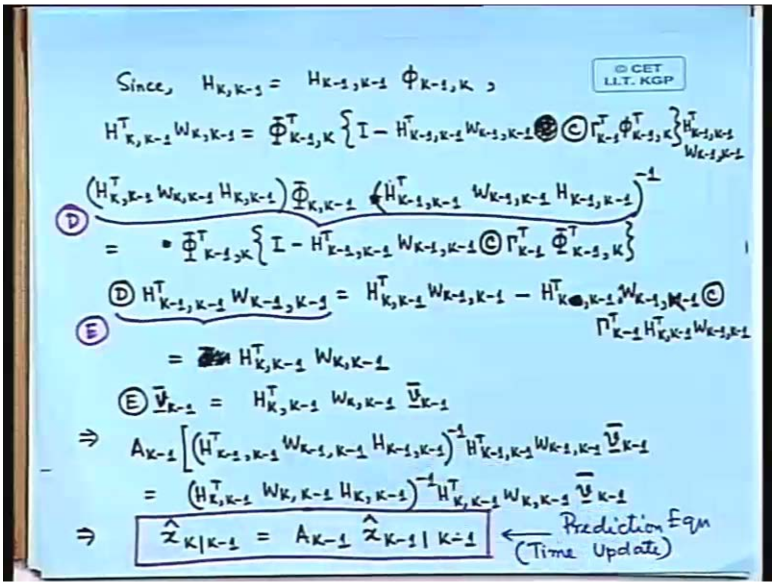
So exactly I one way is to again write these two problems; for now k is equal to k minus one on one side and k is equal to k on the other side, and then try to relate them that is the rigorous approach then written here. But then before we come to that, let us get a sort of you know, what is called as sneak preview kind of staff; that is just imagine you are I mean you are you have some state, right. So you have how is xk plus one and xk related? Ak, xk plus gamma xi k, this is how it is related, correct. Now you have already made an estimate of this; that was your x hat k given k that is the best estimate possible, because you have taken measurements up from zero to k into account. You cannot get future measurements, okay. So you have done all that you could in this.

Now standing there without any measurement, you want to look forward, correct. So how do you look forward? Obviously how does a system look forward? How does the true state get advanced in time? By like this, correct. Now so obviously you will have to use this rule, but then what about this? The question is that, here is something it is had it been simply this Ak xk ,you would have said that my estimate is xk plus one equal to A hat k, Ak x hat k given k. If it was simply this then you would have also applied the same rule; because there is a there is nothing else; now the only problem is that you have something else, you have this one.

So the question is that, can you estimate even a part of this? That is the question but you cannot, why you cannot? Because this xi in this estimate, you have taken care of all process noise up to k minus one and you have already made an assumption; that the process noise is uncorrelated, in the sense that xi k the you cannot get any information about xi k from all from zero to k minus one, that is the that is the nature of the random process.

So so from your previous measurements of xi zero to xi k minus one, you cannot get any information about xi. So if you do not have any information about xi k; what is what is the best strategy? Best strategy is to assume that it is in the mean value, because it is completely random. What is the mean value? Zero. So so still even if this is there under the assumption of that stochastic characterisation, it is still best to just propagate with through A, this is an this is an intuitive picture. Now now the same intuitive is actually you can you you you can rigorously compute, again the those two solutions for k minus one and k and then and then arrive at the same.

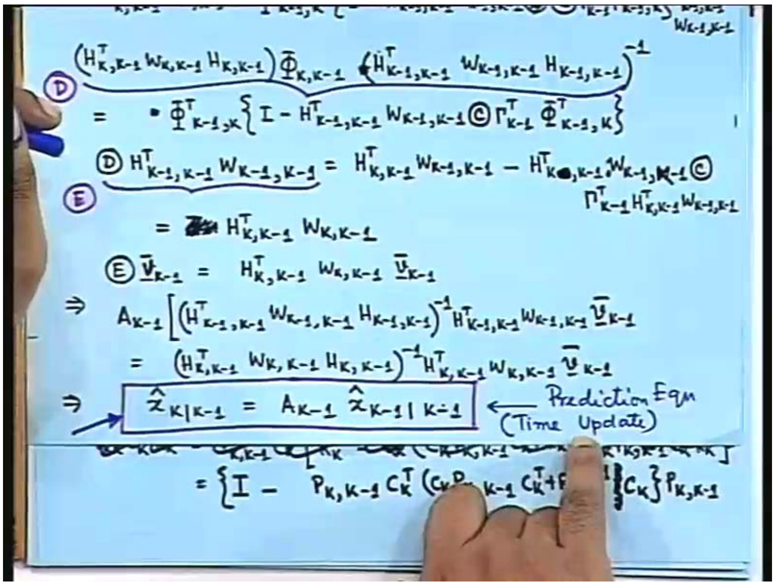
(Refer Slide Time: 42:50)



See all these because of the fact that, that you have to you what is involved is what is what is known as, the matrix inversion lemma which of you might have studied in some other courses A plus, B, C, D inverse is equal to there is there is there is a big formula. A inverse minus A inverse B transpose… I mean I do not recall what the formula is, I always look it up.

So you can again go through that same thing and go on substituting all that, but finally you will get this.

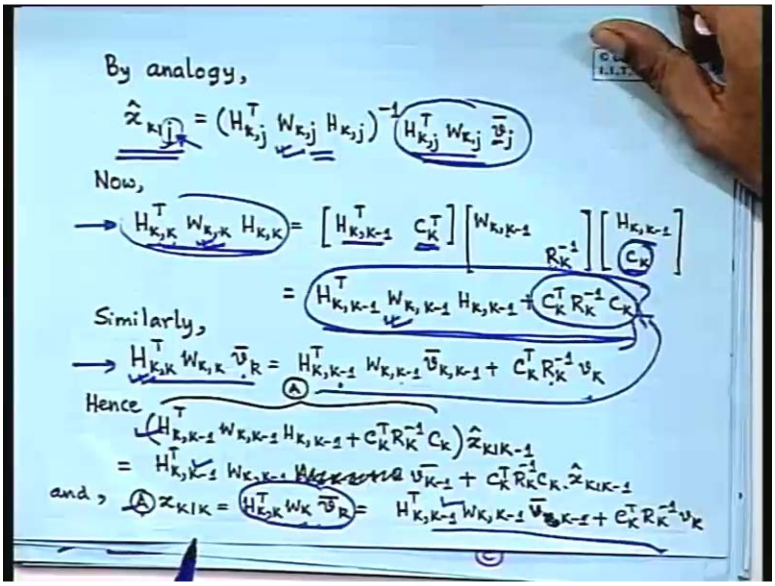
(Refer Slide Time: 43:17)



What we intuitively argued; that if you want to take, here is an estimate of hat of xk minus 1, which you have generated using measurements up to k minus 1. Now you want to propagate the state up to k without having any additional measurement. So if you want to do that, you you you you multiplied simply by Ak minus one. There is no better that, you can do simply because you cannot get any other information about the process noise from the past, that you have seen. So all these mathematics, all these equations we will finally give you the next equation of the Kalman filter; which is called a prediction equation or sometimes called a time update, because here you are advancing time and in other words without any measurement beyond k minus one, you are trying to predict the state at k. Therefore it is called a prediction equation.

So either you have a predictor equip, predictor corrector formulation, there are various terms which are mentioned in the context of the Kalman filter; but these two are the basic state update equations. Now but the Kalman filter has has many other equations, why? It has at least three four other equations, that is because of these, actually these two are the status.. basic state estimation, equation are these two; this prediction equation and this corrector equation. But now people will people wanted to say that this Gk, this Gk also should be you know efficiently computed again recursively; that is every time I do not want to solve, you see what is Gk? Gk is this A inverse, this now what does A involve? A involves big big quantities.

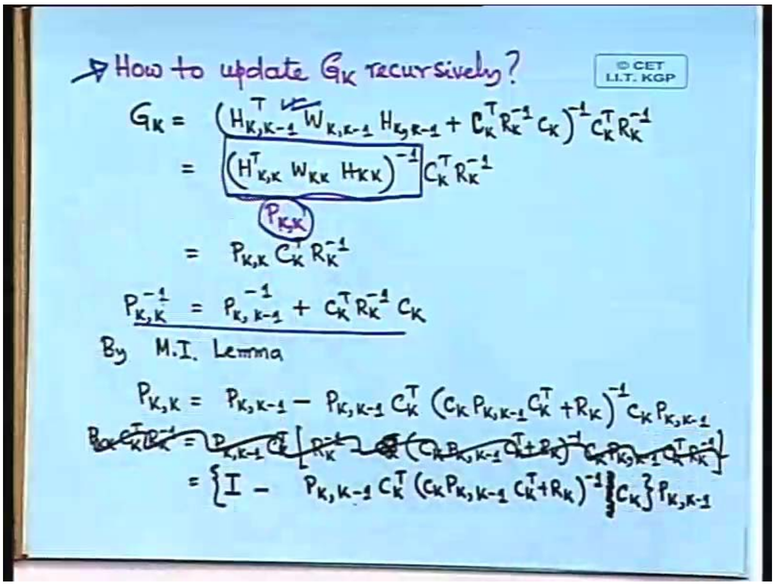
(Refer Slide Time: 45:15)



A involves A, what is A? This one, now you see these again increase in length as k increases. H increases in length, W increases in length; so you cannot compute A inverse using this equation. So therefore again you have to bring that into some constant, constant computation time kind of form, right. So the rest of the equations are rest of the equations of the Kalman filter are devoted just to make this computation of Gk a constant computation which you does not increase with time, because otherwise you cannot use it.

So five minutes, so the question is how to update Gk recursively?

(Refer Slide Time: 46:03)

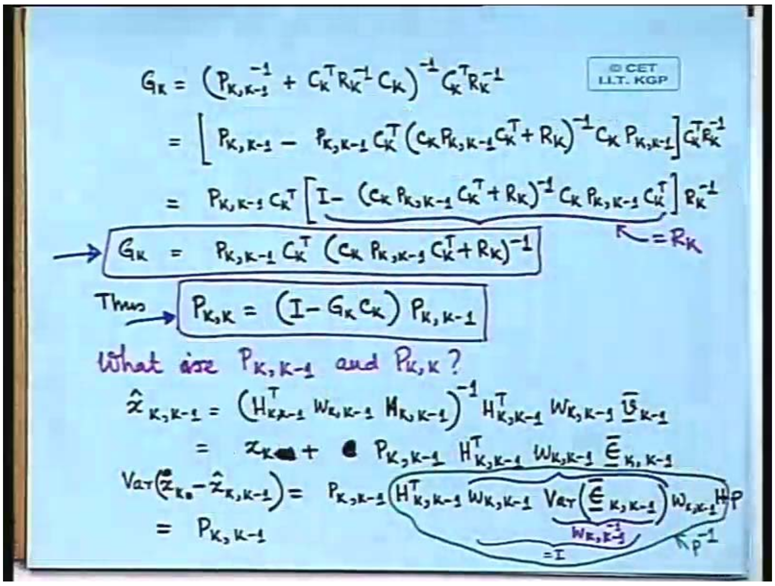


It is given Gk minus one, how to get Gk from there using a constant computation at all k. So now again again a matrix algebra. Now this is this is just I have restated that formula; A inverse this again written it for your reference, this is nothing but this, let us give it a term called, Pk k. This is sometimes called the process noise covariance, no not process noise covariance, state estimate covariance, state error covariance, why that we will see. But let us give it a name, we have given the name Pk k; because here also k, here also k so we have given Pk k. So one thing is clear that, this is and correct. Obviously; if if this is Pk k, this inverse remember that, Pk k is the whole thing including the inverse. So then Pk k inverse is the inner thing, just this one.

So obviously Pk k inverse equal to from this equation itself, Pk k inverse equal to this equation; this this is this follows from here to here. Now there is a matrix inversion lemma, again matrix inversion lemma is arrive in different forms. You know so again matrix, you know this is nothing this the it it is a kind of matrix identity; people have said that, A plus B, C, D inverse if you do, it equates to that. It it it is the identity, it will hold for all matrices which compute; so no problem it is just happens. It just a you know a cute observation, which somebody made regarding matrices.

So so lot of manipulation, putting this here, that here, taking left pre-multiplication ((00:48:05 min)) common and all that; You can you can read it, you can try this, this is not, mean I mean cannot really I mean you know recite these equations here. So finally it turns out that,m you can cast it in the in these two equations; this just nothing but manipulation, if you see in the notes, in fact I have tried to make it more simplified than the source.

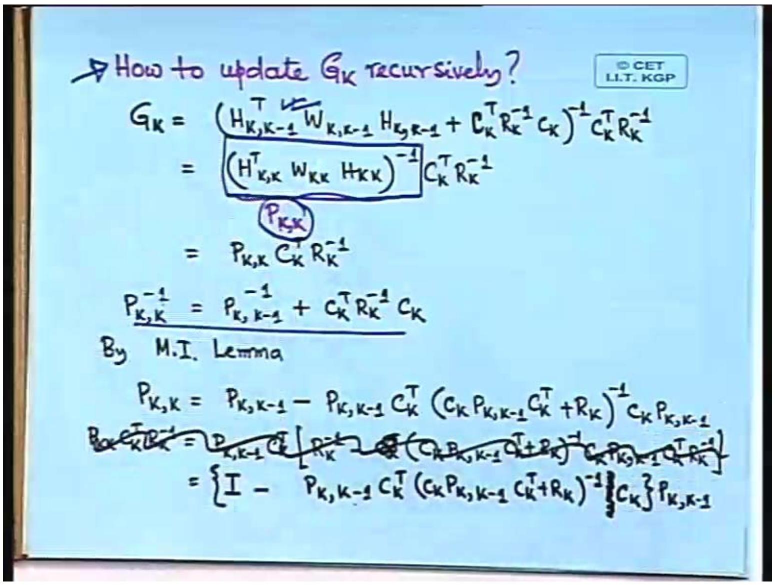
(Refer Slide Time: 48:19)



I mean the the source from which this is adopted, that is chuan chens book; it none of these steps are are actually elaborated, where it will say, it can be shown, it it can be shown, it can be shown, but I have shown for your benefit, spending some midnight twelve.

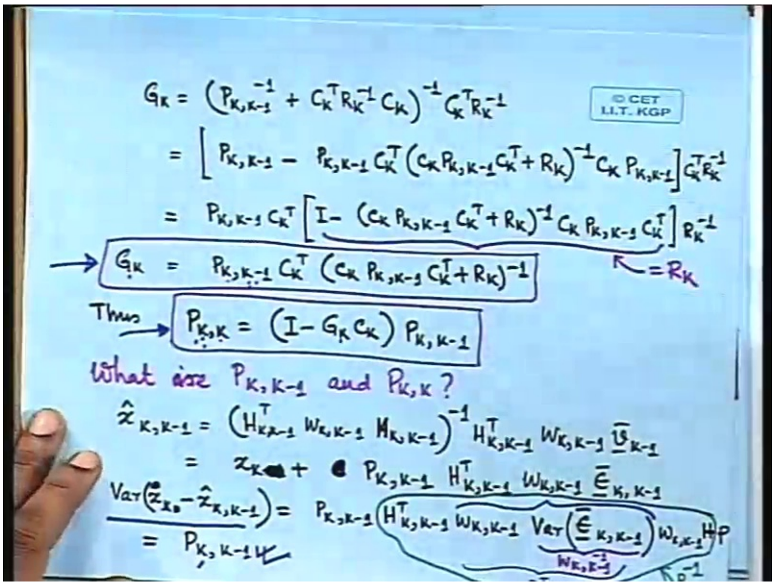
So so finally you get this. This is nice see these are all these are all constant dimension matrices. So at least there is an inverse here which is slightly nagging but but this there is there is nothing whose dimension increases. So you have all constant matrices at at each instant k, same computation. Now there is a as I said that, this Pk's are called state error covariance's, why?

(Refer Slide Time: 49:20)



I mean I just took some arbitrary form, I mean I just puts took some arbitrary formula matrices and then named it, Pk k. So so how I can tell that it is a it is a state error covariance, that has to be proved.

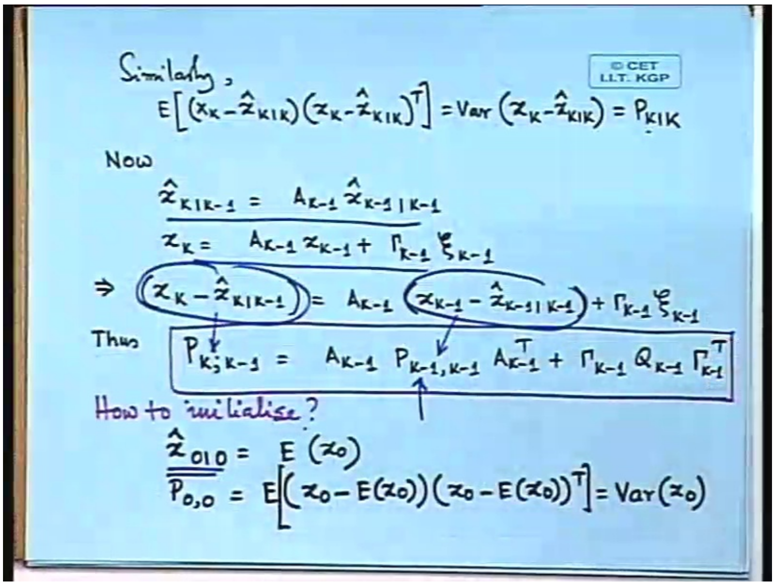
(Refer Slide Time: 49:35)



So you can actually prove you can actually prove that, this proof is also given here; that variance of xk minus x hat k given k minus 1, evaluates to Pk, Pk k minus 1.

So what is this? Thi is state error. If you make an estimate of xk taking measurements up to k minus one and then take the error from the true state; then and if you take it variance, then it will try to it will become Pk given k minus one. So therefore this is the covariance matrix of state error; that is why it is called state error covariance. Similarly by exactly same matrix algebra, you can you can compute even k given k, and you will get Pk k, same thing right.

(Refer Slide Time: 50:30)

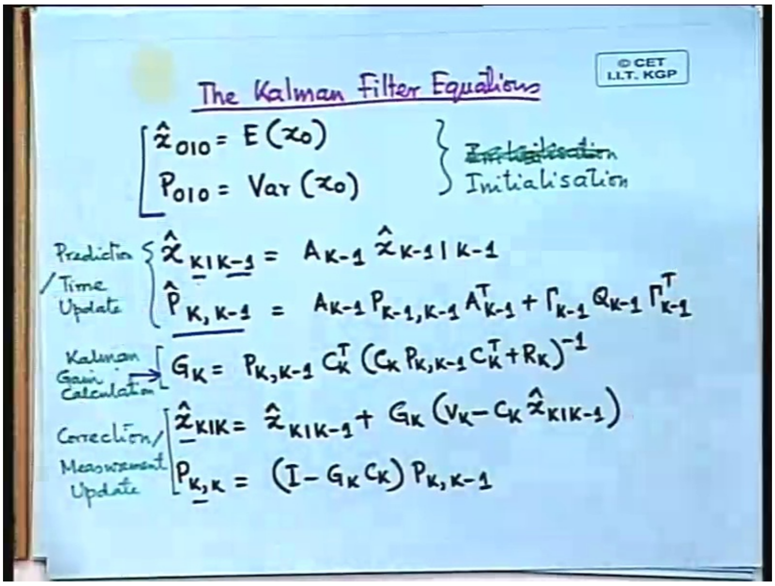


So final thing, one there is there is only one thing remaining, so you see now what what have you what have you done in this equations? In this equation, you have calculated from Pk, k minus 1 to Pk, k. See using Pk, k minus 1 you can calculated Gk, using Gk and Pk, k minus 1 you have calculated Pk, k. So from k, k minus one you have gone to k, k; so again it is like a measurement update, even for this you have to make a time update because everything must proceed in lock step.

So you have to now, final equation is that you have to take k k minus 1, you have to get from k minus 1, k minus 1, you have to update this one. So that is very simple, that is that is because simplify observing this equations. So this two are standard equations; so you now compute xk, k minus 1, this side you get k minus 1, k minus 1. Now if you have a if if you have a vector which is like you know Ax plus B, then the vary say say y is equal to Ax plus B; y is a vector, A is a matrix, x is a vector, v is a v is a vector. If you have such a case then obviously variance of y will be A, variance of x A transpose; this is the this simple. So so just from this equation, you can get this equation because the variance of this is this and the variance of this is this, is very simple. How do you initialise this algorithm? You have to start with something because you are, every time you have to give some basic estimate of x hat zero, given zero and P zero given zero then you can take it up from k to k minus one.

So that is here is this this is your initial guess of x hat, you you know nothing about it. You have you have no measurement so far; so therefore you.. you you have what whatever is your mean estimate of your initial state, that you must put. And then P zero, zero will become the variance of so; that is why I said that you need to make assumptions about, the initial state as a random variable. You you need to make an assumption about its mean, and its variance. So that is how you initialize the algorithm, right. So finally we have made it in time, just in time by skipping some steps these are the Kalman filter equations.

(Refer Slide Time: 53:10)



So you first initialise then after initialisation, what you do? Now you have no measurement, so from there you predict to x one. So from k minus 1 you predict to k so. First you apply the prediction equations, because you have no measurements so far. And similarly you update k to k minus 1 here. So these are your, so these are your time update equations, then you get your first measurement.

When you get your first measure; in the mean time you calculate Gk then after you get your first measurement, you correct this. And then again predict for the second measurement and then get, then again predict for the second time step and then get the second measurement correct it; then then again predict for the third time step. This is the way you go on and you get amazing results as we will see in the next class, thank you very much.