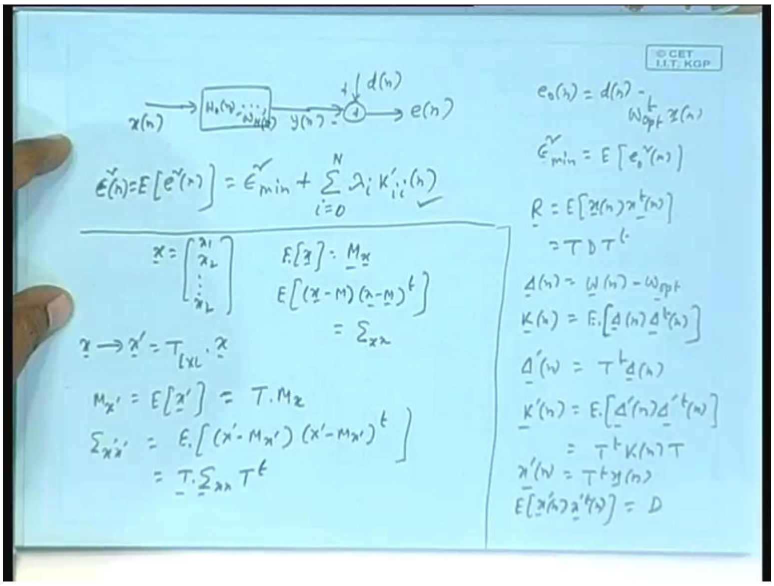
Adaptive Signal Processing

Prof. M Chakraborty
Department of E and EC Engineering
Indian Institute of Technology, Kharagpur

Lecture - 10
Convergence Analysis (Mean Square) – Contd

(Refer Slide Time: 01:02)



So, we come back to our customary drawing with which we begin every lecture you know everyday which will continue for some time till I switch over to another topic. So, I am dealing with real valued case, because you know I was halfway through some analysis. But, last time we are analyzing this error epsilon square m is equal to this and do we found to be the minimum plus an extra component that I am coming to what is the minimum. So, minimum is when that is if you consider e o n as D n minus the optimal filter transpose x n vector there is when you are use an optimal filter.

Then the error is u n epsilon mean, in fact I will write epsilon square to be in confirmative with the notation it is this, there is I want to filters to the ideally the optimal filter. So, that then E n has variance equal to epsilon mean, but unfortunately we do not get that as I told you because in that steepest descent analogy an a derivation, we did not replace R and p while their exact values.

But, where approximate value and that is why the convergence takes place in mean there is weights do not convergent the actual optimal filter, but into dances around that in the steady state. So, obviously the E n in the steady state will not be having variance as slow as epsilon square mean, but will be saving some extra component and that over analysis showed to be something like this. So, I equal to 1 or 0 both possible 0 to n lambda I k prime I n. Now let us define things, redefine I mean just for our recapitulation point of view we had all matrix was an input autocorrelation matrix all input.

But, x n are D n they are 0 mean these are all standard then we had R equal to T D T transpose T consists of ortho normal Eigen vectors D is a corresponding Eigen values. So, R is assume to be positive definite which means Eigen values are not only real they are also positive and T D T transpose T is unitary matrix. So, T transpose is equal to T transpose, T is equal to I, there is T transpose is the inverse of T and vice versa the all that we know. Then we had delta n, our main purpose was to see this I mean we define this delta n, you remember what is the what does delta n mean, it is the deviation of W n from W opt k n was this side.

So, I am keeping a kind of dictionary for all u, usage this is the k n, k n is k n is E. this are all we can say covariance matrix of the tap weight error or tap, error tap error or weight error covariance matrix. But, correlation mat covariance matrix of it was delta you are taking is called is the steady state delta n as 0 mean as you have seen. So, correlation and covariance should be in the same nevertheless, it is called weight error co weight error delta n is weight error this is called weight error covariance matrix.

Then from delta n, we defined for analysis another transform vector which is T transpose delta n, remember this delta n transpose delta prime n is a vector this had a diagonal correlation matrix. So, covariance matrix that is if you that k prime not di, sorry this not diagonal not, this is not diagonal, sorry I made a mistake, this what is k prime n k prime n was. So, if you really replace delta n by this thing here what is k prime n k prime n is e of delta prime n let us delta prime transpose n and if delta prime in if you replace here you will get T transpose k n T there is not diagonal.

But, if you define another thing x prime n as T transpose x n then the components of x prime n will be uncorrelated if you see its correlation matrix. So, that will be diagonal which equal to D that you have seen that is E of x prime n transpose n, it will really replace T transpose x, x transpose T x, x transpose is R T transpose R T, T transpose RT will be d. So, component of x prime n they are uncorrelated they have got a diagonal correlation matrix this is our dictionary of definitions.

Now, this k I prime n is comes here along with i-th Eigen value, so obviously if this quantity this is the quantity at n-th index at any n-th index. So, that it mean square value of the error it is not epsilon square mean, but something more which is this and k I prime, n is the i-th diagonal entry of this transformed covariance matrix at index n multiplied by lambda i. So, with time n we have to make sure that this quantity remains bounded that is if you take the limit of this quantity as a n tends to infinity this should not grow with n.

But, this should be even bounded and the bounds should be our control by some parameter by which we can keep this as low as possible. Then our job will be done epsilon square n can never be epsilon square mean, but then they can be made close to each other at least epsilon square will be within some finite bound. So, finite limit or finite range from epsilon square means this, what we have to do, so that means we need to analyze this is where we stop last time. So, this analysis are very lengthy analysis and you have to carried out, but I am now formed the background by which we can easily appreciate many steps.

Now, only one background I have to form that I will do now suppose forget all this, now suppose I have got a vector x consists of elements say x 1, x 2, dot, dot, dot say x L, L number of random variables may be 0. So, it means random variable you can assume to be like simple and it is said or me you need not assume to be 0, means it is said that they are jointly Gaussian they are jointly Gaussian with some mean m. So, that means if you take expected value of x that is E of x 1, E of x 2 dot, dot, dot, E of x l you form a vector or means at that you call m this collar mean vector and E of x minus m x minus m transpose there is a covariance matrix.

Now, say these are replications this all this derivations definitions I gave in the very beginning of this semester this you can call sigma covariance matrix in the special case a where m equal to 0. So, the 0 mean random variables sigma is nothing but the correlation matrix correlation covariance matrix they are same then suppose this is given. So, you want to transform x vector to another vector x prime linearly that is x prime is some key matrix times x square matrix T, T is L cross L, L cross L times of your choice times x.

Then firstly x prime also will consists of random elements after they are coming from x 1 x 2 up to x L for this if I call it to m x if I call it sigma x, x. So, what is m x prime that is E of x prime tap lie E over this e over this linear combination you can take the matrix out you can directly apply you over this. So, linearly combine the elements will be summation something times x 1 plus something times x 2 got dot dot. Then expectation operation, better you take the expectation on the elements of x and then you will get a same thing.

So, it will be nothing but T into E of x and E of x is m x, so m x prime is T m x and what is sigma x, x prime that will be your x prime minus m x prime x prime minus m x prime transpose. Now, you replace x prime by T x m x prime by T m x and so obviously you can see we will get this to with T times sigma xx T transpose you replace this T x, T m x T taking out here also T. But, T get transposed so comes to right hand side, so irrespective of whether than Gaussian or any other, I mean with respect to their probability density joint probably density this is always true.

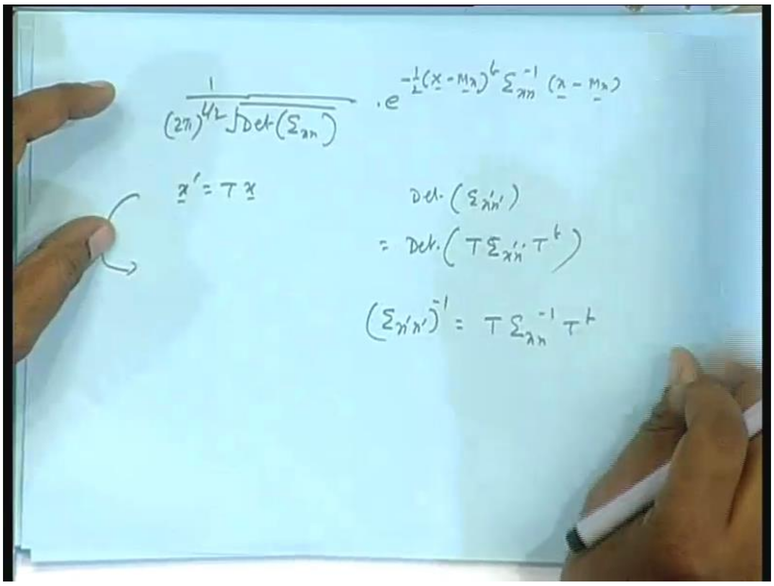
So, that if x is a random vector with this kind of mean then this kind of covariance then after being linearly transform to another vector x prime they are mu mean vector and correlation covariance matrix they are like this. But, these always true irrespective of the probability distribution of covariance joint probably density of this element say x 1 to x l. Now, assume that this to be Gaussian or a set of jointly Gaussian variables linearly transform to yield another set of variables x prime.

Then this real sign I cannot prove here, because this of part, this is if you take a book control it a theory by Papoulis or some other book, we will see these are basic results. But, there is no scope to prove, here some of things you have to accept that in that the Gaussian is so beautiful.

But, you know in the case of Gaussian distribution, joint Gaussian distribution that is if you want to vector x of jointly, I mean random variables which are having a joint

Gaussian density joint probability density which is Gaussian. Then if you linearly transform it, x prime also you consists of a set of random variables which are jointly Gaussian this is the basic result, so if x is consisting of jointly Gaussian variable. So, will be x prime, but you remember that a joint Gaussian density needed only two parameters, one is the mean vector another is a covariance matrix, it is m x n like a.

(Refer Slide Time: 12:28)



So, these are the formula 2 pi to the power l by 2 L, L number of variables were there into determinant of that covariance matrix times square root. But, I think sigma x, x into E to the power minus x minus m x transpose sigma x, x inverse 1 by there is 1 by 2 here these are the definition. Now, my claim is that if x if you have x prime as T x then the x prime, also we will have a similar density function sigma x, x sigma x, x prime they will be replaced by. So, can be obtain easily from sigma x, x by the previous formula or same, here some of the nice thing you can see that if T is a unitary matrix.

So, suppose T is a unitary matrix just as a deviation suppose T is a unitary then that is suppose x if you find out the correlation matrix of x R matrix decompose at T D T transpose T is a unitary matrix. Now, pick up that T, T x is x prime say in that case determinant of here we will have sigma xx prime sigma x, x prime for the mu here.

But, what is delta would you understand I am saying the join T density for the elements of x prime that will go involve which term determinant of sigma x, x prime. But, sigma x, x prime will be what determinant of T to this T inverse at determine this called singularity transformation determinant will be same as the determinant of the original matrix. So, this value and this value will not change because T is unitary that is T transpose is same as T inverse. Here, when the inverse comes inverse of this will be what you will have another term n a sigma x prime x prime inverse that will be what replace sigma x, x prime by this take the inverse.

So, T transpose sigma will come first T transpose inverse is T this will come, I mean I do not need this further analysis it is just something I am telling you also that of. But, the basic result is there that if you have a set of Gaussian distributed vector a, if you have a vector consisting of jointly Gaussian random variables. Then if you linearly transform them, the resulting elements also are jointly Gaussian that is a basic result it can be proved. But, then it will concern some of the theory and all that I have no time that proof is not difficult to read, actually if you see any probability books say the book by Papoulis you see.

So, you first start with transformation of random variable that is given a random variable x with some kind of density if you now consider a function F of x calling it y. So, what will be the probabilities of y as a function of y not as a function of x, as a function of y a then they will generalized it to nut not from x to y. But, a set of variable x 1 to x l to the set of variable y 1 to y l and they are it is not difficult is given in books you have to read just a well theorems and results. But, the Gaussian things will comes back as Gaussian fine, now for this analysis, now I have to basically see the behavior of this guys with time this is my target.

Now, I have to see the behavior of this guy with time which means I have to stickle this mat is better I start with this matrix, after all this matrix what are these elements there are the diagonal entries of this matrix i-th diagonal entry is this of this matrix is this. So, that let me rather try to track this guy how this guy behaves with time as n tends to infinity from that I will find out how its diagonal elements behave with time right. So, let me instruct a tackling the diagonal elements let me rather tackle the entire matrix take care of the unitary matrix.

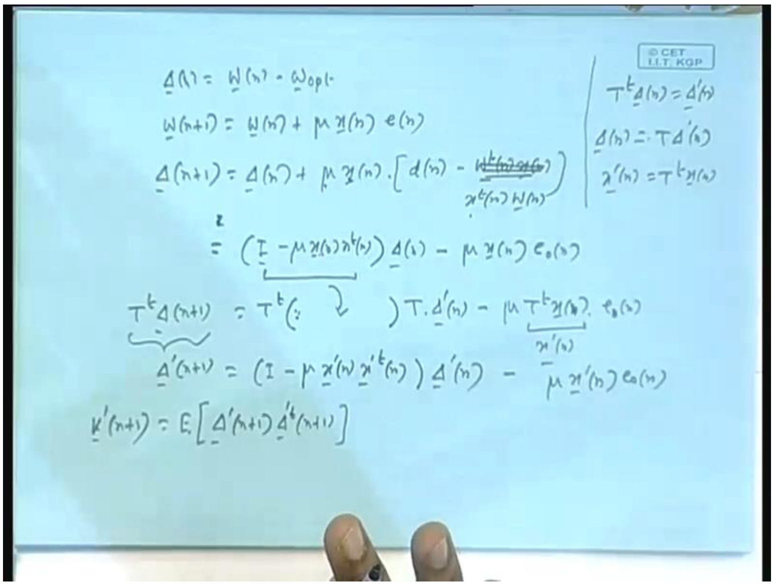
Now, let me see how it behaves with time as n tends to infinity if that be I do not know once again you have a mode of analysis is what like in the convergence of the elements. So, what you have seen we first define delta n then found a recursive equation of delta n, delta n plus 1 in terms of delta n and that give raise to some conditional mu under which down square of delta inverse tending to 0. Here also I will try to develop a recursive equation in what u matrix, however the diagonal elements of the matrix.

Now, that is what will be the values of these elements at n plus 1-th index giving their values at n-th index and then I will starring the recursion and see how it goes as time tends to infinity then we may approach. But, it will take many steps, but many of we can then work on fast, because we have given the background, for that analysis I will make one assumption. But, another assumption not only independents assumption independents assumption is always there for this analysis also that is current weight vector W n is statistical independent of D n and x n vector.

So, that is always there we will be further assuming that D n and the elements of x n, they form for at least elements of the vector x n, they form a set of jointly Gaussian variables jointly Gaussian. Now, I will assume probably they are all random, but there is some correlation between D n and x n that is why I am trying to estimate it, and we will make that assumption that they are joint density joint. Now, density between whom D n and the components of the x n vector that is x n x n minus 1 x n minus 2 dot, dot, dot, x n minus capital n plus 1 or x n minus N.

Now, that is a jointly Gaussian vector the as a vector consists of those variable which are jointly Gaussian, this is I assumption, we will make and which a, this a fine assumption. But, Gaussian assumption usually works in practice this is the next assumption anyway, so let us, as I told you I want to track this matrix k prime n. So, what is k prime n, it is the covariance of this guy or correlation of this guy rather delta prime n, now what is delta prime n T transpose delta n, so let us see what is delta n.

(Refer Slide Time: 19:09)



We know delta n was actual weight minus the deviation and we knew W n plus 1 is by the elements this part we have already done. Now, I am dealing with purely real valued case extension to the complex cases exist, but there is why much more complicated no point. So, if you subtract both sides if you double, if you subtract W opt from both side you will obviously get delta n plus 1 as delta n plus mu x n E n and E n is D n minus W transpose n x n. So, this part is common you know we have done this analysis isn’t it this part is E n W n you replace W n as W opt plus delta here.

Then if you take the W opt transpose x n D n minus that that will be the optimal error e o n, so see if you replace W n by this plus this there are two terms coming out of it. So, one is W opt transpose x n D n minus that that will be the e o n another is minus delta transpose x n, but instead of W transpose x n permit me to write it in the other way. So, x transpose n W n both are same under real case at least both are same, so W n I will replace by W of plus W delta n, x transpose W opt D n minus that, that will be your e o n and there is another term x transpose n delta n.

So, I have got delta n then mu x n x transpose n delta n, can I skip in the intermediate step you can see possibly, identity matrix delta n I into delta n will give you this I into delta n that is taken care of. So, there is another delta n coming here x transpose n delta n and before that mu x n, so mu x n x transpose n delta n x.

So, x transpose is a matrix at minus sign, so minus mu x n x transpose n there is that matrix that times delta n coming out of dull del W l and the other term is if you put W opt here D n minus this part is the e o n. So, mu x n e o n, but as I told you I am not interested in this here I have go here we have got k ii prime n, k I prime, n comes from k prime n. But, k prime n comes from delta prime n delta prime n comes from T transpose delta n, where T transpose came out of the factorization of R this is called factorization of R factorization of R.

Now, I mean like you know there are three factor matrices multiplied give you this, so this T transpose, so that means delta n is not enough I would take T transpose delta n. So, find out this then the correlation of that I have to do for n plus 1-th index and write that entire thing in terms of it is value at n at the n-th index. So, that means delta n plus n plus 1 is not enough, I have to multiply this by T transpose this is my other part, so I have to multiply here also. Now, I defined T transpose delta n is my delta prime n, remember delta prime is the quantity of interest, here I may take the covariance matrix of that.

But, the i-th diagonal entry that is why you forget delta n get into delta prime n, but if you have this, that means what is delta n T transpose you take it of the other this side inverse of that inverse is T only. So, T transpose is unitary, T transpose in a, so here I am multiplying from left hand side by T transpose, so T transpose will come here this thing this entire thing come here. But, delta n, I will write as T into delta prime n minus again mu I multiplied from left by T transpose, so I have to multiply this also by T transpose and e o n is a scalar is remains as it is.

Now, this quantity by my definition is x prime n remember I define delta prime and also x prime, x prime n was T transpose x n and what is the correlate covariance matrix of x prime that was diagonal d. Now, by T transpose, I develop two vector, delta prime n and x prime n delta prime n from delta n x prime n from x n delta prime n because I am interested in the covariance matrix of delta prime n not delta n. So, because the matrix that was coming into operation there was delta prime k prime n here or, therefore here not k.

But k prime k prime comes from delta prime that is why I am getting into delta prime and delta prime is this, but similarly I am using this also T transpose x n and this as diagonal correlation matrix D I coming from R Eigen values of R. Now, you see T transpose first element here is I and then T, T transpose I T is T, I only you are you replace this here the first term is I T transpose I into T. So, T transpose T is at the i, so that is equal to I only and T transpose, forget the mu T transpose x n that is x prime then and x transpose n T that is x transpose T that is same as this.

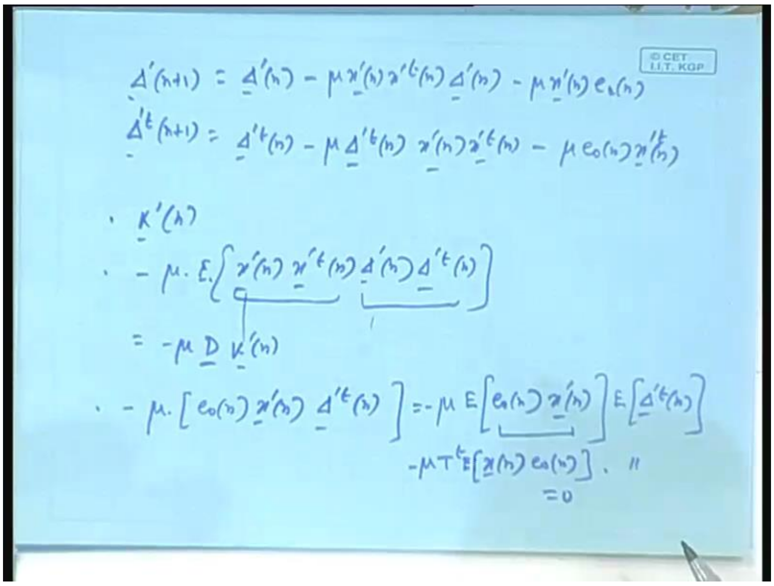
Now, if you take the transpose of this you are going to get x transpose T, so that time delta prime n minus mu this is as it is e o n. Now, I want to find out k prime n plus 1 that is covariance of this mind you this is very lengthy analysis very lengthy we have not done anything here. Now, so far, but this will give you some idea about how to carry out this analogy, how this to make some assumption to make life simpler this is that we will have good excise. Here, what I told you in the beginning that I will study the matrix k prime n in a recursive manner k prime n plus 1 will be evaluated as a, in a recursive manner in terms of k prime n and others.

But, that is why I form delta prime n plus 1 in terms of delta prime n and something else then I am finding the auto, I mean the covariance of that this into its transpose. So, that is k prime n plus 1, here I will replace delta prime n plus 1 by this entire thing here also, so you understand how big the product will be 3 terms, 3 terms, 9 terms. So, I will have to analysis that is why it is very messy, but many things will become 0 by our clever manipulations, but you understand even nine terms I have to handle, so this is very important.

So, I will consider all the nine terms 1 by 1, first is remember delta prime n plus, please see this you should not make any mistake you know there I say chance of making silly mistake. Here, delta prime n plus 1, here that means no transposition on this, but then next 1 is delta prime with transpose. So, this and the same thing, but every term will be transpose and you have to multiply for your convince you can write this and again write the same thing here. But, put a transpose on everybody and then do like, you know 2 multi polynomial multiplications like that, but I will not go that level go down to that level.

So, here one term, first term, first term will be what E is always there delta I into delta prime n that is delta prime n below delta prime transpose, please get use to this when I kept transpose of this. So, I into delta prime n there is delta prime n only that transpose if you are not understanding, let me write down the two terms.

(Refer Slide Time: 28:50)



This will only become instead of putting in bracket for you convince I am throughing out of bracket, so I into delta prime n that will become minus mu this into this. So, this one terms and another why you want delta prime transpose, we will use transpose I am doing like a you know the way things are done in school take this take its transpose. So, if you need a transpose of this if you need a transpose of this is a vector, this is matrix.

So, this will come first and then transpose of this, but this is symmetric curve you see any vector into its transpose is a symmetric matrix, any vector any vector x, x transpose y, y transpose z, z transpose there are all symmetric matrix hermitian matrices. So, transpose is same always you can verify take the transpose of this comes, first transpose cancels like that. So, it remains as it is and the other 1 is mu because this is scalar you can write e o n times this transpose or this transpose e o n either way same, sorry there is a transpose here. Now, we multiply the 2 whatever we doing after all this vector into its transpose and then e over that, so this vector into transpose e over that.

Now, let us work like school boy this into this, this into this, this into this then this into this into this, this into this then this into this there like that 9 terms. Now, you understand that first we do this into transpose and that will be what see the first term will be I mean this k prime n only. But, there are plenty other terms then next is this, let me is this visible this blue color visible because other day there was some problem that I can multiply x actually. Then next 1 is minus mu e this part into this, but I assumed what is delta n delta n is W n minus W opt and x prime n is obtained from x n directly T transpose times x.

So, I assumed in that independent assumption that W n is independent of x n and therefore delta n is independent of x n. Therefore, delta n is independent of x prime n also that the x prime n is just obtained from x n nothing else. So, that means this e o n matrix did you understand it in the previous class if you really this is a matrix this is a matrix. But, if you really multiply 2 matrices apply E of T you can unscramble, you can separate down the terms involving x prime separate terms involving this. So, instead of doing that at the end you can as well apply E over this, E over this and then multiply you will get a same result.

So, can you I think you can foresee that , so if you do that now, so these are the thing I have made you prepared actually, so that is why these things will not take much time. So, this will be nothing but E over this part, E over this part, E over this part is we will have covariance matrix of x prime n, but there is a diagonal matrix D mu D and E for this is nothing but k prime n. Then the other one is mu x prime e o n this and then you form it I can write e o n in beginning then x, x prime n then this transpose. Now, you understand x prime n is independent statistically independent with delta prime e o n that also is independent because e o n consist of what D n minus W opt times again x n vector.

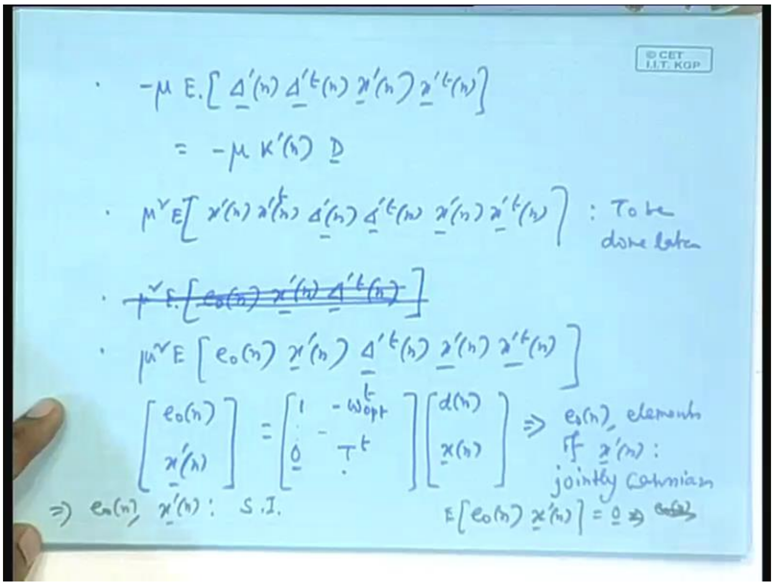
So, it depends on D n and x n vector by my assumption delta n and, therefore delta prime n there are independent of D n and they are follow delta n and, therefore delta prime n because delta prime n of a purely obtainable from n and vice versa. So, delta n therefore delta prime n by assumption independent assumption, there are independent of what x prime n x n. Therefore, and x prime n and e o n because e o n depends on D n and x prime again x n again, so that means you can separate out this part e over this e over this.

Now, what is this quantity e o n is the optimal error and x prime n is T transpose x n, so this quantity you can write as E, you can write T transpose out minus mu T transpose. So, you can take out e o n is a scalar only this is T transpose x n, e o n you can push on the right hand side T transpose x n into e o n. Now, e o n is a scalar, T transpose you take out times this quantity, but e o n is orthogonal to all the components of x n because e o n is the optimal error. So, correlation is 0, so this will give rise to 0 please understand it is not purpose only to present this thing in the adaptive filter context.

Now, you will learn this term and how to carry out some statistical analysis how to make some clever assumption here and there, this is the main purpose of this course. Now, it is not a course on probably statistic, but it is a course which will give you some training for statistical analysis of signal and system that is the main. So, that will be a main game of this course, for this course not to study the just adaptive filter this is for is 0, so 3 gone, I have left with only 6. So, next will be so this fellow is gone next is this guy delta prime n with this with the, is this with this.

So, this again delta prime and delta prime transpose x prime x prime transpose E over that you can apply E over this part separately this part separately a very much like this. Here, x prime is first side delta prime on second side then the other way, so I write the next term because I have, I am shifting this page. But, you have to trust me I will just draw I mean from there only, you can check in your with your notes also.

(Refer Slide Time: 36:49)



So, then next term is minus mu obviously you will, you can work on this part, you can work on this part this part will give you diagonal matrix T this will give you k prime n. So, this is minus mu k prime n diagonal D, so far so good this term this is huge term, the biggest, this we will have to do separately this again need a special this into this I am writing it. So, this is mu square minus minus plus mu square E, it is so difficult to remember x prime x prime transpose delta prime n delta prime transpose n, again this thing comes back x prime n x prime transpose.

So, not difficult to remember only thing is because too much of space to it we did it in this quantity to be done later, to be done later, I will do it separately it needs some other steps and all that and the last one is between these two. So, again minus minus plus will be mu square mu, mu square, now this will be 0 obviously you can see we have already done it somewhere here same thing. So, sorry this we did it, this will not come this and this I have already taken care I also by mistake I was writing it back again. So, that is wrong this with itself this 1 mu square to sorry see how tricky it is even I am getting confused.

Here, again it is very difficult to analyze that is where Gaussian assumption will come, but then I have told you 1 thing that if a set of variables are uncorrelated. But, it is not guaranteed that there are statistically independent, but if there are statically independent that always uncorrelated. But, if one case, only one mean the other and vice versa that is when the density is Gaussian, a set of Gaussian random variables if uncorrelated mean then also statistically independent and vice versa. So, remember that another thing here you see delta prime n and x prime n by our independent assumption they are statistically independent.

But, this and these are statistically independently, because x prime n is obtained from x n and this depends only on W n, so there this, 2 are independent. Similarly, e o n and delta prime transpose there are statistically independently, e o n consist of D n and x n and it consist of delta n W n, so independent. But, I do not have only this if I had only this much like before I could have just separated out and by using orthogonal realm this could be 0, I am getting x prime n back again x prime n consist of x n e o n depends on x n.

So, these 2 are not at this there is some relation between them and then x prime n comes back here comes back and all that, so I cannot apply that here. So, again I will apply this again this is all you know I mean tricks I will apply that Gaussian thing, now you see e o n, can I write this way e o n and this vector. So, if I form a joint thing W opt is row vector column vector W opt transpose is a row vector minus sign these are not given in books by the way book again copied from somebody else. So, these details I have not been spoken about in the books, first see the first row 1 into D n minus, this is a row vector W of transpose x n minus of that.

So, that will be a mu n no problem with that and then I have got 0, 0, 0, 0 vector in the first column and then this sub matrix is T transpose. Then can I not get this 0 every 0 here will be take care of D n, D n will be D n will have no effect we will multiplied by 0 from this column of zeros. So, we left with T transpose time x n that will give you x prime n by our definition, so these vector of how many x prime x prime consist of n plus 1 and one more n plus 2 number of random variables.

So, N plus 2 number of random variables is obtained from another vector of n plus 2 number of random variables by a linear transformation this a matrix only. But, I assumed these elements to be Gaussian, jointly Gaussian that was third assumption I made independently. Now, assumption 1, 2 and then third assumption that the D n and elements of x n vectors they are jointly Gaussian and from these vector by linear transformation I am getting another vector. So, that mean there are also jointly Gaussian, this implies e o n and elements of jointly Gaussian.

But, we have also seen we have also seen because of orthogonally that optimal error we have also seen that e can you see this in the screen E, e o n times x prime n all the components. So, if you take e o n and multiply with each component take expected valve you get 0 because after all x prime n was T transpose x n, T transpose you can take out e o n is orthogonal to all the components of x n. So, it is 0, so now you see I have got this situation where e o n and the other elements of x prime n there are jointly Gaussian. But, correlation between e o n and the elements of x prime n that is 0 that means there are statistically independent Gaussian.

So, uncorrelated e o n correlation with all elements of x prime n, e o n with the first element e o n with the second element e o n in the third element each correlation is 0. But, there are jointly Gaussian, that means they are statistically independent you remarkable, how we did that, we took that joint Gaussian formula. Then the cross correlation terms in that matrix sigma was put to be 0 then the entire density was separable as a product of individual densities sensible work. Here, that e o n, the term from coming from e o n that can be separated out 1 density term will come out, can you see what I am saying.

So, you will have a joint density I have a correlation matrix in the overall joint density that will have correlation between e o n and the elements of x prime n and also there individual correlations. But, those terms will be 0, e o n with the correlation terms involving e o n and terms from the x prime vector. So, 1 probability density involving e o n that can be separated out, so that is why just statistically independents will come up. But, anyway we have proved this result we have to just code this take this result as it is that there are jointly Gaussian e o n and the element they are from the jointly Gaussian set of variables.

But, in that e o n is uncorrelated with the rest, so that means e o n is statistically independently with the rest. So, that means e o n let me write separately here this means e o n and x prime n they are statistically independent if, so that will help me in doing all being this. So, e o n is statistically independent of x prime n and of this guy with these guy because of independent assumption and of these guy this fact, so then e o n can be separated out from the rest. So, this entire e, business e over e o n into e over the rest, now e o n what is the mean value of that 0, because x n is 0 mean D n is 0 mean.

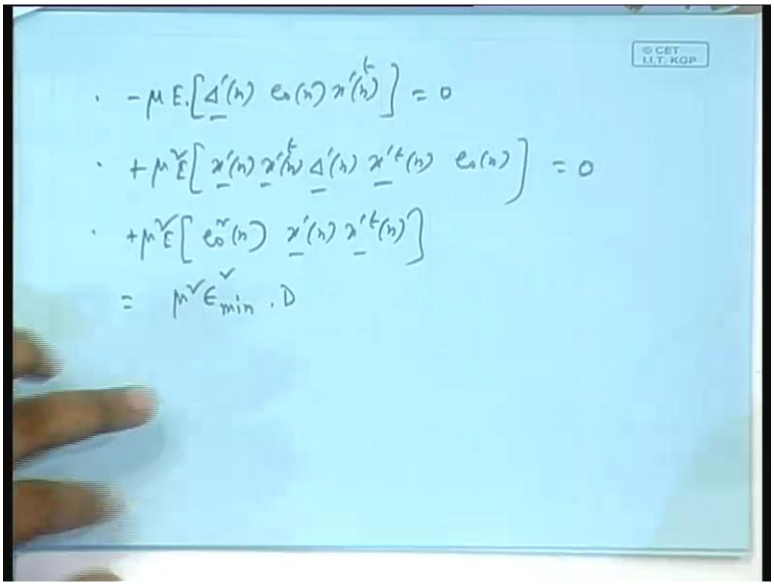
So, D n minus W of transpose times x n that is 0 mean input x n is 0 mean D n 0 mean is it, so D n minus W of transpose x n vector if you take it and apply E over it will be 0, E of D n is 0, E of x n 0, do you follow this. So, that means this will be equal to 0 because I am not writing that steps, shall I, do I write the steps separately that this E is nothing but E of e o n times, that is this times this separately. So, this part is 0 this part is equal to 0 this becomes e over this and e over this, now how because e o n is statistically independent of x prime n also.

So, by independent assumption its already I independent of delta prime n, so e o n can be separated out and expected value of e o n is 0 because e o n what after all D n minus W of transpose x n vector. So, if you apply expected value over it E of D n is 0, D n is 0 mean E of x n vector is 0 because x n 0 mean right, so it is 0 mean, so it means this is 0. So, understand to I mean, I mean eliminate this terms we can you know bringing this some Gaussian distribution and some assumptions. But, all that cleared out, these are things I mean you are getting a training actually you know I mean to invoke as and when required.

So, the Gaussian assumes are works very well in practice most of it, this term is 0, so what was this term, this was this cross middle term into middle term that is the biggest one and this crossed out. So, I am left with only this guy with this, this, this, so 3, what terms that very easily done delta prime n and this that will be 0, can you see delta prime n independent of e o n and this. So, an E over this e o n is orthogonal to the components of this x prime n and that will be 0, I will right down, but I am telling you this we have to do separately.

So, this have to separately here x prime, x prime transpose delta prime x prime transpose e o n separate, so e o n is there, x prime n is there, again x prime n is there again x prime n is there delta prime there. So, again I will use this assumption that e o n itself is independent of x prime n also delta prime n also. So, e o n part can be separated out e of e o n will come separately which is 0, are you following that same thing, this part I will here also where when I handle this product this two. So, this two regions of e o n, e o n is scalar e o n, you write the right hand side and rest x prime x prime transpose this, they are 1 hand e o n is independent of that. So, what is e o n can separated out expect the value of e o n that is 0, so that term will go I am left with only this, so I am write it down the terms quickly.

(Refer Slide Time: 50:17)



Next class I will not redo this, so you will to collect this in staff room I will take this pages with me and, so another term is minus mu E of delta prime n this is 0 obviously. Then because this you get separate out with this and this because some orthogonality it will go I am repeating what I stated and that other term is plus, because minus minus plus mu square E. But, here you assume that here you make the, I mean make use of the fact that this fellow is independent of the rest. So, E over this will come out separately which is 0, so this will go to 0 only thing is last term is important.

So, last important is product of this two minus minus plus mu square x prime square there is x prime n x prime n transpose and e o n, e o n there is scalar e 0 square. So, plus mu square E of E 0 square n and e o n, x prime n there are statistical independent we have seen. Therefore, e o n square is statistically independent of this there is the advantage of statistical independence, if x and y independent then x square y cube F x g y there are independent.

So, that means this part is independent of this part, so you can separate out expected valve of e o n square n is the epsilon square mean that is the minimum variance at enable, because these are optimal error. So, mu square because e o n when you have put the optimal filter the corresponding variance is the minimal enabled this and E of this, E of this is your D.

So, in the next class I start from here remember we did not work out this particular term, what term I left us to be done later where is that page this one to be done later. So, this needs some special analysis not very big, but again some fact Gaussian distributed variables. So, that mean a if two or four variables are there x 1, x 2, x 3, x 4, and there is a product x 1 into x 2, x 3 into x 4 and take E over that another Gaussian case. So, you can simplify it at E of x 1 x 2 into E of x 2 x 3, x 3 x 4 plus E of x 1 x 3 into E of x 2 x 4 plus E of x 1 x 4 into E of x 2 x 3. So, just kind permutation you know that fact as to used there to simplify it, so I will do that in the next class.

Thank you very much.