Introduction To Adaptive Signal Processing

Prof. Mrityunjoy Chakraborty

Department of Electronics and Electrical Communication Engineering

Indian Institute of Technology, Kharagpur

Lecture No # 25

Second Order Analysis of LMS Algorithm (Contd.)

Last class, I mean we considered a adaptive fit. But though we did not use LMS or any mode of adaptation just an adaptive filter where filter of coefficients are varying with time and at nth index nth clock filter coefficient vector is Wn and we had Vn was the error between the actual filter coefficient vector at n th index minus the optimal one ok.

$$\underline{v}(n) = \underline{w}(n) - \underline{w}_{opt}$$

And then if Wn is used just for recalling if Wn is used as a filter. So filter output was W transpose sorry filter output was and corresponding error was En is Dn minus this is Yn.

$$y(n) = \underline{w}^t(n)\underline{x}(n)$$

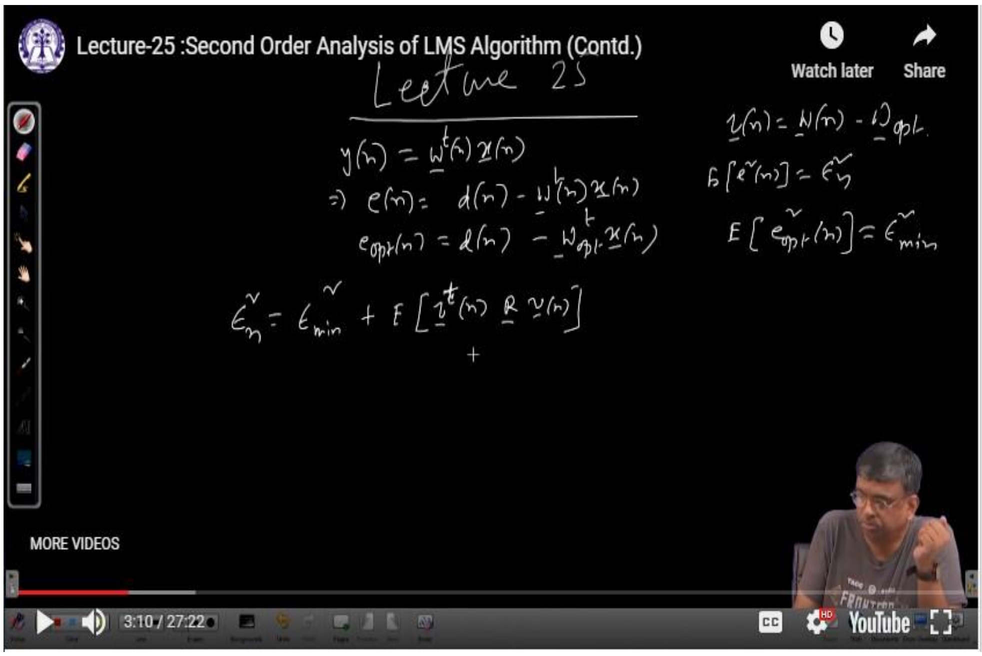
$$e(n) = d(n) - \underline{w}^t(n)\underline{x}(n)$$

This I am not using W opt here this error will have variance power higher than that of the case when I use W opt Dn minus W opt n Xn ok. And that error is called E opt n where Dn depends as it is this is all from previous this lectures a quick recap ok.

$$e_{opt}(n) = d(n) - \underline{w}_{opt}^t\underline{x}(n)$$

So, variance of this is larger than the variance of this because this is why the I use the optimal filter and the difference between them we are trying to work out. And what we found is this that is if you denote the variance of this noise this error as a minimum possible variance. So, epsilon square min and here if you say as just epsilon square n then epsilon square n under some you know independence generalized independence assumption which

I did last time this is having this component plus an additional component sorry R is positive definite and therefore, any Vhn as long as Vn is not 0 ok. So, this will be for any Vn actually Vhn I am writing I could as well as transpose because they are all real I am considering only real case.



So, any V transpose and R Vn will be non-negative and if Vn is positive it will be positive since Wn is different from W opt here will be in general non-zero maybe in one particular trial could be 0, but in general non-zero. So, when you average you get a positive term ok. This is an extra component this is called excess n square error epsilon square excess at n, excess n square error. So, let us work on it we know that R is a positive definite matrix means R is a Hermitian matrix. We have studied that length in the beginning of this course some properties of Hermitian matrices and positive definite matrices.

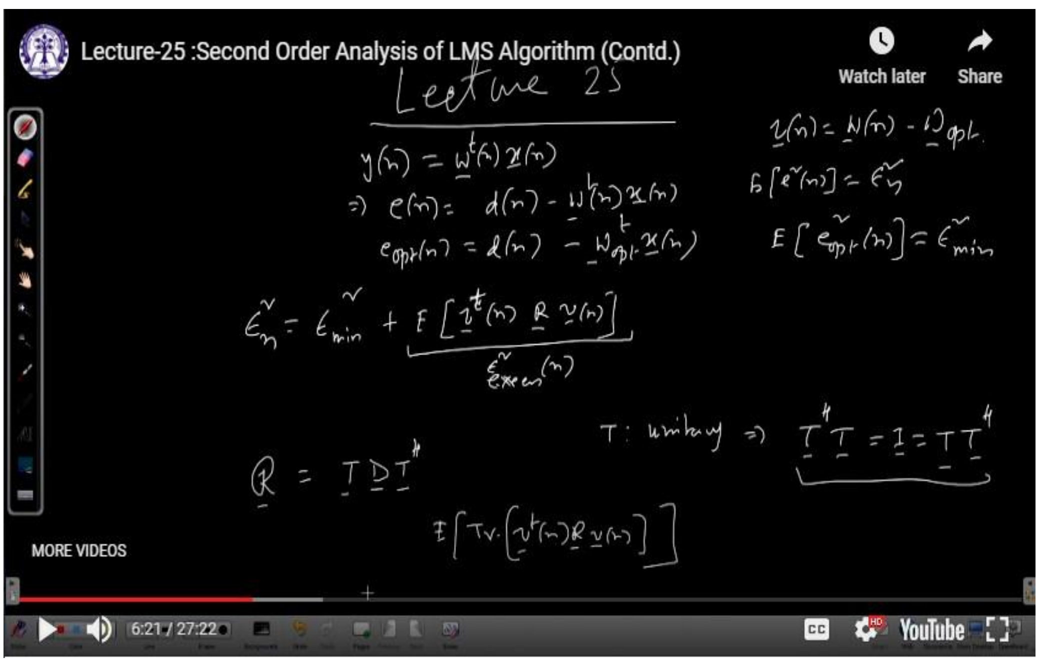
We know in these cases the Eigen vectors of R corresponding to different Eigen values they are all mutually orthonormal that is each having norm 1 and they are orthogonal and if I put them side by side this matrix is T R can be written as T then a diagonal matrix D where D consist of diagonal entries only and those are the Eigen values and Eigen values are positive not only real positive because it is a positive definite matrix ok. And the same

T H, T unitary because all the columns of the T are mutually orthogonal orthonormal rather meaning we have all studied this. So, no question of going back there again T H T is identity which is also that means one is the inverse of the other therefore, you can put T in the beginning and T H here that also is identity ok. This is what we studied earlier all right. Let us consider this expected value of V transpose N or VN I will replace R by this after a while.

$$\underline{R} = \underline{T}\,\underline{D}\,\underline{T}^H$$

$$\underline{T}^H\underline{T} = \underline{I} = \underline{T}\,\underline{T}^H$$

But before that let us see one thing this is a scalar ok row vector matrix column vector matrix into column is a column vector row into column is a scalar right of course, because left hand side is a scalar expected value of that. So, scalar is a 1 by 1 matrix. So, trace we know the term trace, trace of a square matrix is a summation of the diagonal entries. So, a scalar is a 1 by 1 that is square matrix. So, its trace will be that element itself is not it.

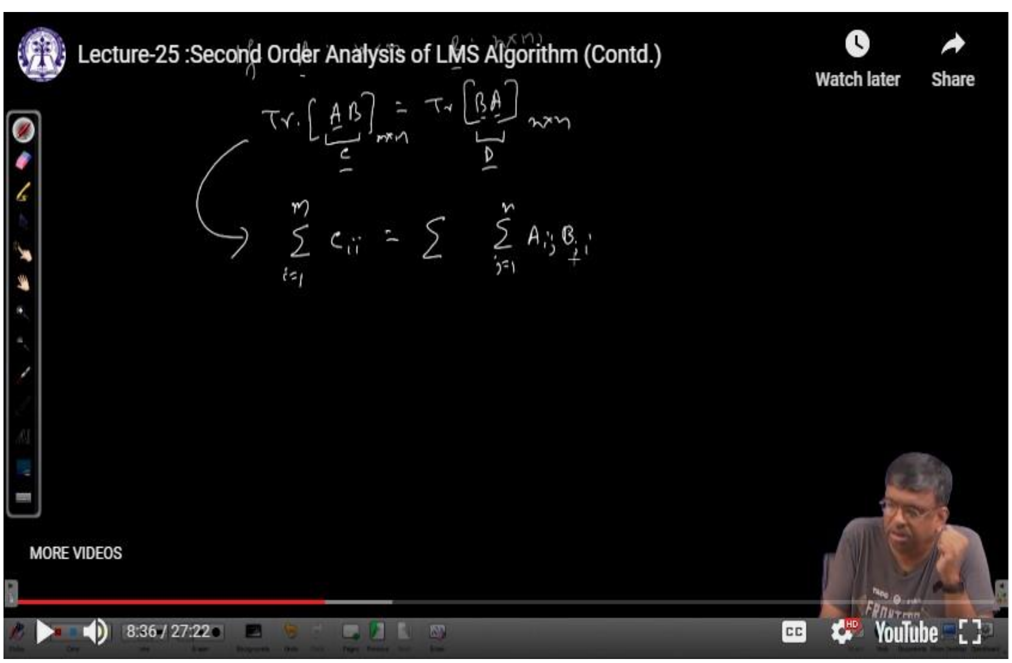


So, this scalar I can write it as a trace this term excess term I can write as trace of this quantity scalar term all right. So, this whole thing is a scalar, scalar is a 1 by 1 matrix which is square and therefore, trace of that is that element itself. So, this whole element is nothing, but its trace. Now there is a result that if A suppose is a m cross n matrix and B n cross m then trace of AB and AB is a m cross m matrix the trace of that square matrix is same as trace of BA which is a n cross n cross n matrix. I am not sure whether you know it or not, but it is very easy to derive which I do for your sake.

$$Tr.\left[\underline{A}\ \underline{B}\right] = Tr.[\underline{B}\ \underline{A}]$$

So, trace of from here we see trace of AB, AB if I call a matrix C and BA if I call a matrix D trace of C means all the diagonal elements are to be added. So, C i comma i is a ith diagonal entry. So, I start at i equal to 1 to how many it is a m-by-m matrix right m cross n n cross m. So, it is a m by m and this is a n cross n. So, i equal to 1 to m.

Now what is Cii? AB you are doing product right. So, ith row of A ith column of B when you are multiplying you are getting Cii. That means A ith row and you span the rows by using a index j and j can be 1, 2 how many A up to n. So, A1 I have not completed yet, but I span like this scan like this i1, i2, i3 like i1, i2, i3, i4 like that. So, I am going along the row and simultaneously I am moving B in this direction j i.



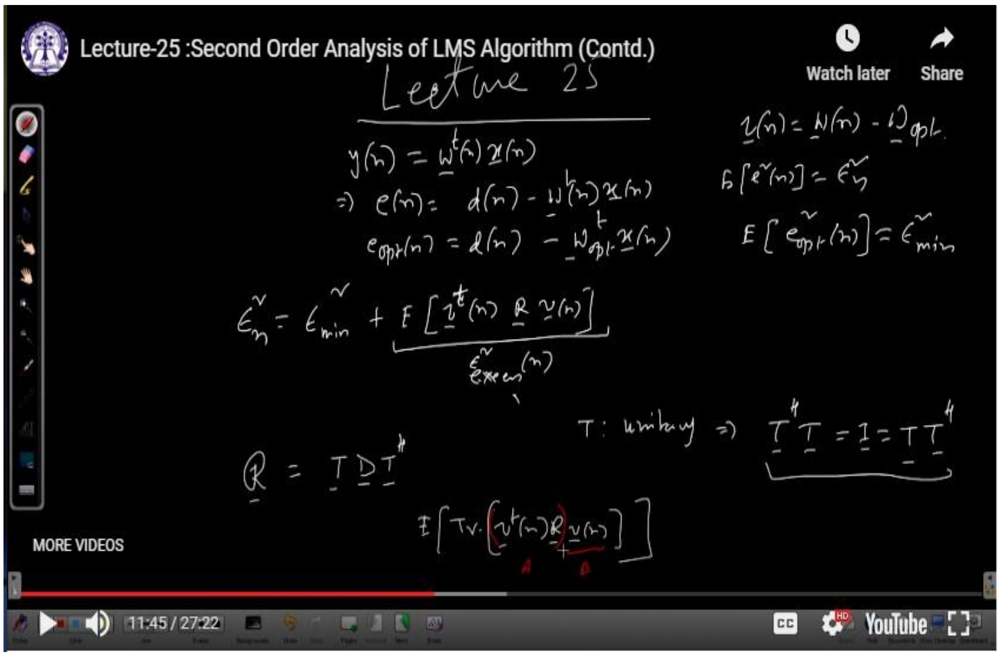
So, ith column so B1 i, 2 i, 3 i like that. So, BA i1, B1 i, A i2, B2 i, A i3, B3 i like that. So, ith row and j ith column ith row and ith column of ith row of A ith column of B they are multiplied that gives me the i comma i diagonal entry of C. So, this is my Cii and then whole thing is sum because of the stress I am taking place over this. Now you know any double summation means what first you fix a value of i for that i you move j over the entire range carry out the summation.

Then again take another i fix that i here and then move j over the entire range carry out the summation and do on doing it and then add all of them. I can do the same thing if I hold j first at 1 take all values of i then again another value of j take all values of i and so on in the end I will get the same result this called interchanging the two summation or summation order. So, first you fixed j equal to 1 to N then you for each j you move i equal to 1 to M A ij B ji I write this way B ji A ij what is the meaning jth row ith column. So, j1 j2 j3 as i moves and here jth column as i moves 1 j2 j3 j like that. So, it is the jth row of B and jth column of A they are multiplied which will give me the j comma jth diagonal entry of this product matrix alright.

So, it will be D D j comma j alright. So, it is nothing but BA if you do and you find out the jth diagonal entry then jth row of B. So, here j is fixed i moving j1 j2 j3 and jth column of A. So, here j fixed and i moving they are multiplied and summed. So, that gives you the j comma jth element of D jth diagonal entry and then you are summing.

So, you are taking a trace of this product trace of BA this proves trace of AB same as trace of BA. Therefore, if you consider this down you take this to be row into matrix is a row that is also a matrix after all M plus 1 ok. You call this to be like your A and this you take to be B. The trace of AB because A and B you know we are not square they are arbitrary M cross N, N cross B. So, M can be something and N can be 1 like here.

So, it is a row vector and so on here. So, trace of AB will be same as trace of BA. So, trace of VN V transpose N R. So, we will have trace of E outside trace of we had remember V transpose R V and this V will come in the front R. Now, you see this is a matrix VN is a column vector V transpose in a row vector.

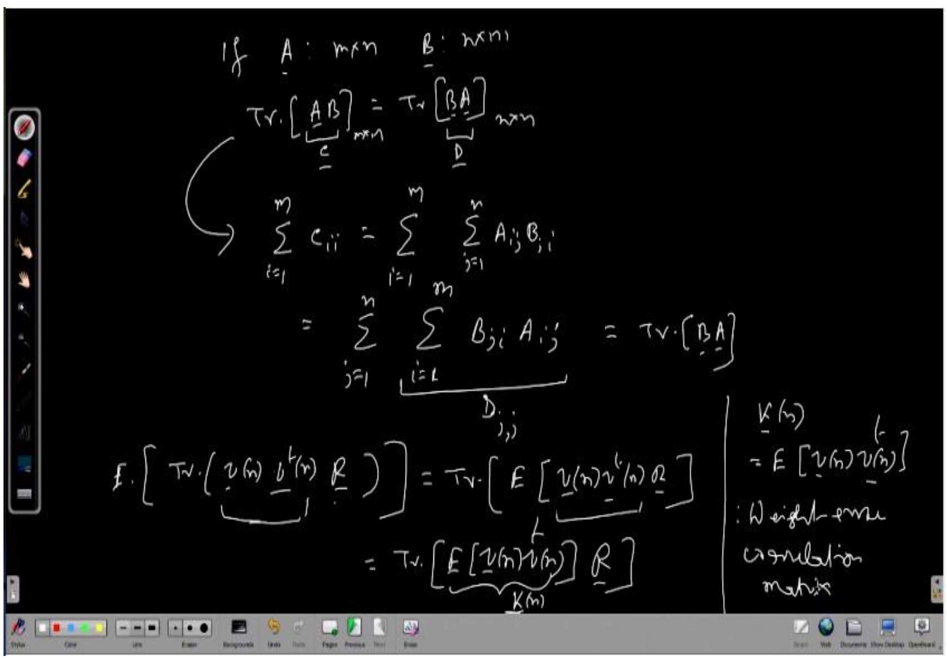


So, this is a matrix this is a matrix this is a matrix together is a matrix, but square matrix Re square VN V transpose is square, square into square is square. So, trace of square matrix then expected value that means, what are with the matrix I take the summation of the diagonal entries then apply E over them. I will get the same thing if I apply E over all the entries of this matrix product matrix and then take summation of the diagonal entries only is it not. Here you first take the sum of the diagonal entries then apply E on the diagonal entries, but you can do this get the same thing if you apply E on all the entries of this total matrix and then take summation of the diagonal ones. We will get the same thing that is called interchanging the two operations.

So, trace E goes in E VN V transpose N R. But R is not random. So, this is a matrix where all the elements are random because VN consist of VN is what WN minus W of and WN is random. So, VN is random. So, this is a matrix we for every element is random that matrix with a constant matrix if you carry out the product and then apply E you will get the same thing if you apply E on every element of this matrix and then carry out the product this we have done earlier.

So, it will be nothing but trace of E working on this part R outside all right. What is VN VN transpose E is the autocorrelation matrix of VN. So, let me call this KN matrix. Remember it is not stationary in general and therefore, I am putting an index N unlike R it was a input autocorrelation matrix independent of N because input was WSS same does not apply here for WN where WN is not in general WSS.

So, I call it KN. KN is E is called weight error vector. So, weight error VN is weight error vector weight error or correlation matrix you can say autocorrelation or simply correlation weight error correlation matrix all right.



So, this I take to the next page these are the next term trace of KN R and now I replace R by what I told earlier Td Th and again take this to be one matrix and this is another matrix AB. So, trace AB is same as trace BA. All right these are thing this I give a new name K prime N which is transformed weight error correlation matrix I will explain why call it like this.

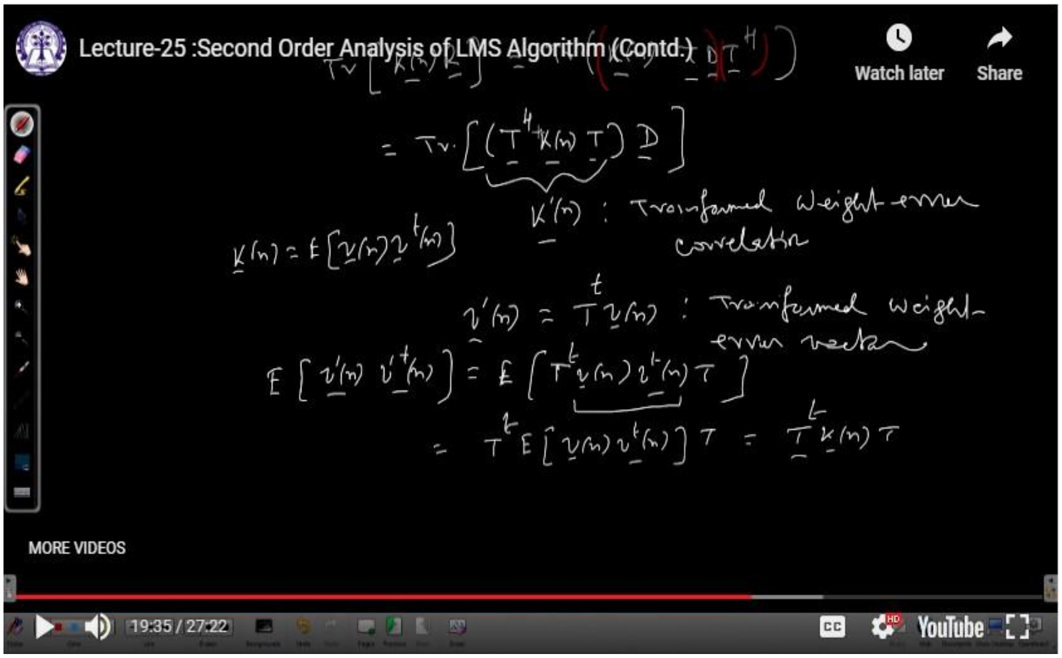
I know KN we have seen already it is E of VN V transpose N ok. Now this T is a Hermitian T is a unitary matrix. Suppose I define V prime N as T working on VN. So, it is transformed T is a transform it works on VN gives you V prime N. So, it is called transform weight error vector.

$$\underline{k}(n) = E[\underline{v}(n)\underline{v}^t(n)]$$

$$\underline{v}'(n) = T\,\underline{v}(n)$$

So, autocorrelation matrix of these V prime N V prime transpose N will be what you replace it TVN here and TVN transpose it will be V transpose N. So, V prime N TVN and V transpose T transpose. So, this is one matrix T is not random constant T transpose not random. So, if we carry out this matrix times this matrix and then this and apply E on that you will get the same thing if you apply because that E will work only on the elements coming from here. You will get the same thing if you apply E on this matrix first and then do the product ok.

This kind of things you have done many times. So, this should not be point of confusion. So, it will be so and this is my KN. So, this is T KN T transpose ok. Let me do one thing instead of this V prime N let me make it T transpose VN.

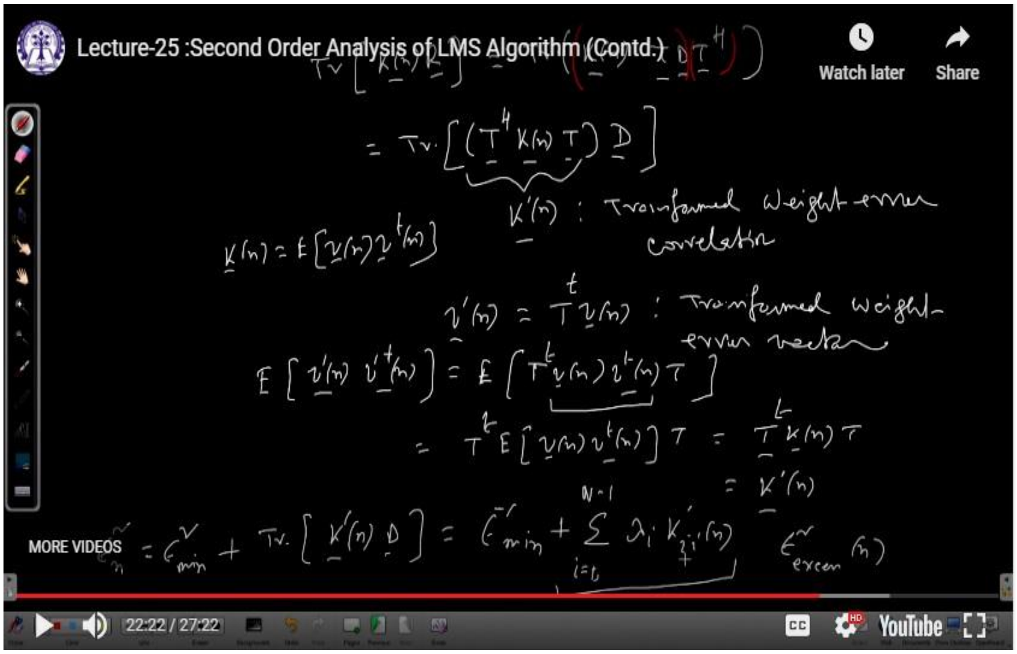


So, it will be T transpose and this transpose will go I am changing the definition ok. So, T transpose and this will go. So, this T transpose and this will go which is my TH here, but we are dealing with real cases. So, T transpose KN T T transpose KN T this K prime N. So, what does it mean this K prime N which was T transpose KN T which I call K prime N and name this transform weighted error correlation matrix.

Actually, if I define a new weighted error vector transformed weighted error vector T transpose VN T transpose is a transform matrix. Then if I take the autocorrelation matrix of V prime N I get back this K prime N. So, that is why it is the autocorrelation matrix of the transformed weight vector it is called transform weight error correlation matrix alright. So, this if we take a look, we have trace KN R will give here ok. So, trace K prime N D and I explained what is K prime N I defined.

So, then combine everything total epsilon square N is epsilon square mean plus this extra term, extra term is this trace of one autocorrelation matrix times a diagonal matrix. Diagonal matrix consists of lambda 1 lambda 2 up to say lambda maybe capital N if there are N such entries ok. So, if you have a matrix square matrix and multiply by a diagonal matrix first column all the elements get multiplied by lambda 1, second column all the elements multiplied by lambda 2, third column all the elements of K prime N get multiplied by lambda 3 so on and so forth. You can verify that then if you take the trace, it will be lambda 1 times 1 1 element plus lambda 2 times 2 2 element plus lambda 3 times 3 3 element and so on and so forth. So, it will be lambda I K prime II N I can be I do not know what range I to call here 0 to N minus 1 or 1 to N or maybe 0 to N minus 1 ok.

We have the diagonal entries this is the extra term ok. This is called excess epsilon square, excess at N alright.

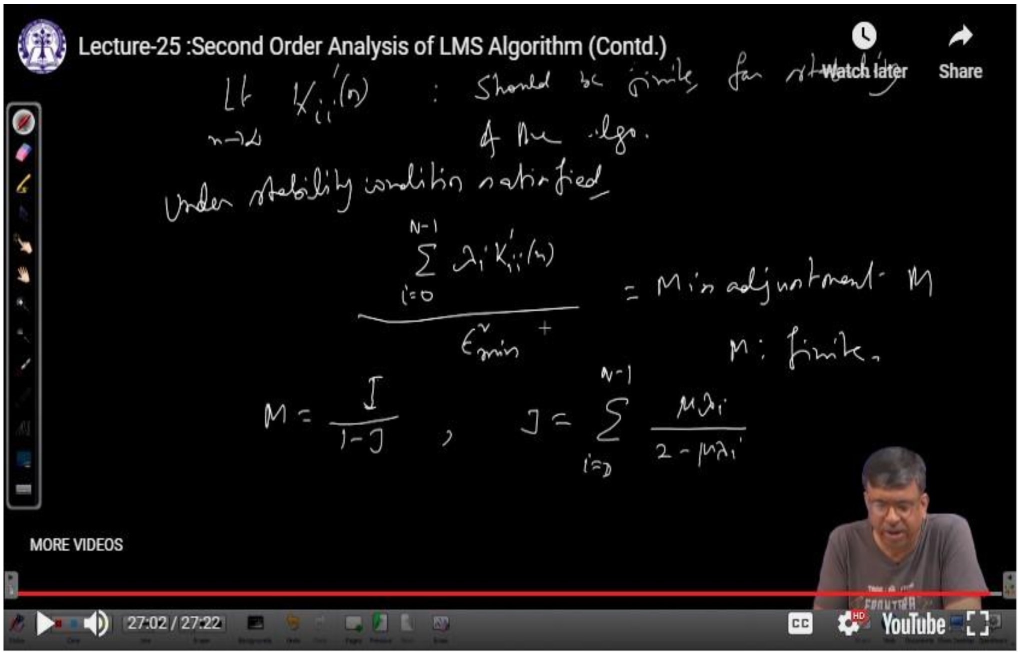


This is something that it is a superposition of capital N number of sequences. What sequence this K prime IIN this sequence. If each sequence does not grow with time, but decays or at least remains at a finite value that is as N is to infinity, they have a finite limit then the overall sum also will have a finite limit and these are all non-negative because they are variances in auto-correlation matrix I am taking the diagonal terms and lambdas also are positive because they are coming from capital R matrix which is positive definite.

So, these are actually positive sequences positive valued sequences. If every sequence goes to a finite limit as N tends to infinity that is in the steady state, then the summation also will lead to a finite limit it will not grow, but if it grows and grows if any of the mode if any of them each is a called each is called a mode convergence mode. So, there are capital N convergence mode each mode is characterized or given by one particular K prime IIN sequence ok. So, if any of the modes now becomes unstable goes to infinity suits up then the same happens for the whole and therefore, this error becomes unbounded and that is a thing we want to avoid at any cost because the error goes unbounded then the algorithm fails is not it. So, therefore, we should make sure that these functions actually each of them converges to a finite limit as N tends to infinity ok, they convert this to finite limit and then we will try to see how much that limit is on what does it depend and how to make it small ok that will do our purpose ok.

So, that means, what we will do is we will evaluate this should be for stability of the algorithm. Now, under stability condition therefore, this is satisfied under this if so happens then obviously, this entire thing is finite ok. We just normalize it to epsilon square this mean epsilon square mean and give it a name this term summation lambda i as it is i equal to 0 to N minus 1 just divide by normalize to this is called miss adjustment. Obviously, under stability condition and we denote it by m miss adjustment m should be finite ok. Now, this m can be worked out this is a very lengthy procedure though very educative procedure one get to gets to learn lot of techniques tricks and all those, but then in this short course I cannot go through the derivation.

So, I will give you the formula of m that we will derive and that is m equal to 1 j by 1 minus j where j is our summation mu lambda i for each eigenvalue i equal to 0 i equal to 1 up to this. So, I have one term mu lambda i by 2 minus mu lambda i. So, mu lambda 0 by 2 minus mu lambda 0 plus mu lambda 1 by 2 minus mu lambda 1 dot dot dot dot ok. This is j and j by 1 minus j is m ok. Remember m is what a positive number because it is variance excess means square error by minimum mean square error.



So, it is a variance positive thing. Therefore, m will consider only when it is positive and then analyze is dependence on j and all that we will find out when m will not suit up to infinity because if it is unstable m will suit up to infinity all those things we will do in the next class. Thank you very much.