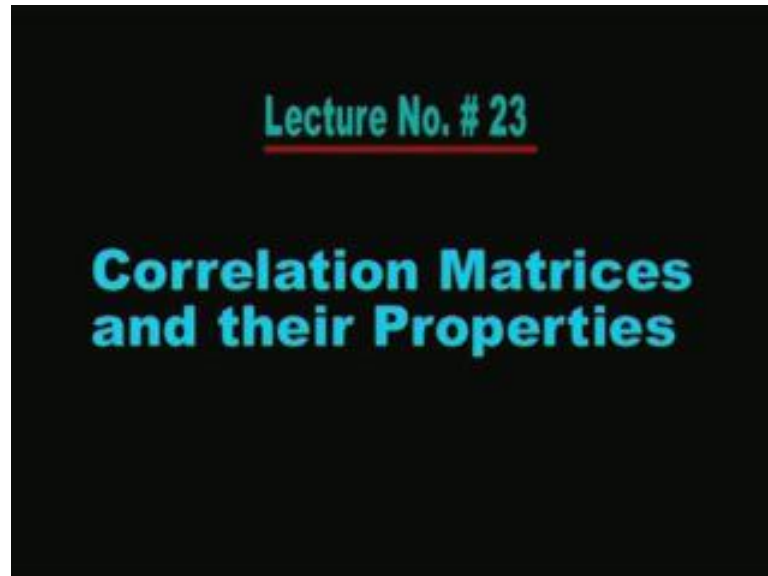


Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 23
Correlation Matrices and their Properties

(Refer Slide Time: 00:41)



So, in the last class, we have been discussing these correlation matrices and in that connection, I talked about what is called hermitian matrices and hermitian transposition and things like that.

(Refer Slide Time: 01:10)

$a_{ij}^* = a_{ji}$
 $A^H = (A^*)^t = (A^t)^*$
 if $A = A^H \Rightarrow A$: a Hermitian matrix
 $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$, $A^H = \begin{bmatrix} a_{11}^* & a_{21}^* & \dots & a_{n1}^* \\ a_{12}^* & a_{22}^* & \dots & a_{n2}^* \\ \dots & \dots & \dots & \dots \\ a_{1n}^* & a_{2n}^* & \dots & a_{nn}^* \end{bmatrix}$

So, today we continue from there. So, just to recall what we did last time, given a matrix say A , we defined Hermitian transposition that is A^H which actually is nothing but A conjugate. Take the complex conjugate of each entry of A and their transpose which is also equivalent to doing it the other way. First take the transposition of A , and then a complex conjugate. That is very simple.

So, if A is not a matrix, but just a vector say column vector, then its hermitian transposition is what we first transpose it. So, it becomes a row vector, and then takes the conjugate of each element. Similarly, if A is a row vector, there is hermitian transposition will be a column vector with all the original elements are complex conjugated. If it happens if A is a hermitian, then A is called a hermitian matrix meaning suppose A is like this one. They are all complex entries A_{11} , A_{12} , A_{1n} and what is A^H . You take the transpose, and then conjugate. So, A_{11} , A_{21}^* dot dot dot A_{n1}^* , then A_{12}^* . Of course, it is A_{11}^* A_{22}^* .

See the elements; diagonal elements do not change their positions upon transposition. Then, A_{21}^* dot dot dot, A_{1n}^* , so A_{1n} comes here. Actually on transposition i , j th element that is i th row and j th column elements become j th element, that is j th column and i th row. Now, if it happens that is as matrix that A and its hermitian transposition, they are same. Then, we will call we will say that A is the hermitian matrix. That means, A_{11} and A_{11}^* should be same, that is similarly A_{22} , A_{22}^* . That means that

the diagonal entries are real and A_{12}, A_{21}, A_{12} is A_{12} star, that is A_{12} in the matrix A itself.

The ij th element and ji th element, that is $1, 2$ th element, that is $1, 2$ and $2, 1$ th element A_{21} . They are complex conjugate of each other. If that happens that is a_{ij} and a_{ji} , they are same in magnitude and reverse in face that is a_{ij} . A_{ij} is a_{ji} star. This is the key relation. Then, the matrix A will be called a hermitian matrix. Now, hermitian matrices play a very important role in this. All A in a probability statistics stochastic process at a signal processing operation is based on this stochastic process concept, statistic concept. So, this matrix satisfies many properties and we will investigate some of these properties. We will discuss them.

(Refer Slide Time: 05:38)

The image shows a handwritten derivation on a whiteboard. The equations are as follows:

$$\begin{aligned} (AB)^H &= [(AB)^t]^* \\ &= [B^t A^t]^* = (B^t)^* (A^t)^* \\ &= B^H A^H \end{aligned}$$

Now, in the previous lecture we have seen that if you have got two matrices which can be vectors also and there is a product AB , and then the product matrix is A transposed using the hermitian transposition, then this is nothing but take the conjugate of the hermitian product matrix. First take the transposition, and then conjugate and then we know AB transpose is nothing but B transpose A transpose conjugate, and you know conjugate of a product of all these reality matrix will consist of terms which are products of terms elements from A and B or B and B transpose, and A transpose and in the conjugate on each product is nothing but product of the conjugates.

So, if you do that, it becomes B transpose conjugate times A transpose conjugate and B transpose conjugate or B conjugate transpose. They are same which is nothing but B hermitian, and then A hermitian. So, transposition hermitian transposition also follows a similar role on product matrices that ABH is same as BH AH, all right. Why we are interested in hermitian matrices?

(Refer Slide Time: 07:00)

Suppose, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$: a random vector

$$\mu = E[x] = \begin{bmatrix} E(x_1) \\ \vdots \\ E(x_n) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}$$

Suppose a random vector meaning actually we had a sequence of n elements x_1 which is a random variable followed by the elements x_2 which is a random variable dot dot dot. Last element is x_n which is again a random variable. We form a vector using this x . Now, first we define the means vector μ vector, that is expected value of x . What is expected value of x ? It is E of x_1 dot dot dot E of x_2 x n. That is equal to μ_1 μ_2 dot dot dot μ_n .

(Refer Slide Time: 08:11)

The image shows a whiteboard with handwritten mathematical expressions. At the top, it says $R = E[x x^H]$. Below that, it says $[R: \text{correlation matrix}]$. Then, it shows $= E \left[\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} [x_1^* x_2^* \dots x_n^*] \right]$. Finally, it shows the resulting matrix $= \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{bmatrix}$.

Then, with these definitions we define a matrix R . As I told you last time whenever I use A , I have a matrix. I will denote it by an upper case letter like this capital R with an underscore. Whenever I have a vector, I will denote it by a lower case letter with an underscore underline, and when no underline is present irrespective of whether I use a capital letter or a lower case letter, it will be a scalar. Suppose I define R as $E[x x^H]$, what does it mean? If the R is called correlation matrix, what does it mean? E of that is 1×1 vector $x_1 \times 2$ dot dot dot upto x_n , and there is another vector upon hermitian transposition we have got $x_1^* \times 2^* \dots x_n^*$. So, that means you have got such elements you can call it R_{11} . I will tell you what is the definition, what is the meaning of capital R_{11} R_{12} dot dot dot R_{1n} .

(Refer Slide Time: 10:26)

$$R_{ij} = E[x_i x_j^*]$$

$$R_{ji} = E[x_j x_i^*] > R_{ij}^* = R_{ji}$$

$$R = E \left[\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} [x_1^* \ x_2^* \ \dots \ x_n^*] \right]$$

$$= \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{bmatrix}$$

Then, $R_{21}, R_{22}, \dots, R_{2n}$, then $R_{n1}, R_{n2}, \dots, R_{nn}$, where this is your R . I write here, where R_{ij} is actually the correlation between i th and j th random variable defined like this as I told you in one of my earlier lectures in the recent lectures rather. I mean we can define the correlation between complex valued random variables like this. E of $x_i x_j^*$ obviously R_{ij} and what is R_{ji} . R_{ji} is $x_j x_i^*$. Quite clearly if you take a conjugate of this, you get this relation R_{ij} and R_{ji} , they are conjugate of each other.

So, R_{12}, R_{21}^*, R_{12} and R_{21}, R_{12}^* is the conjugate of the other, that is you can say R_{12} is equal to R_{21}^* , R_{21} is equal to R_{12}^* and like that. This also shows that R_{11}, R_{22} upto R_{nn} , they are all real numbers obviously, because if R_{11}, R_{11}^* , they are same that is a number and its conjugate are same that be the number has to be real. You can see it here also what is R_{ii} , that is E of $x_i x_i^*$ which is E of mod of x_i square and mod of x_i is always real value. So, it is E of mod x_i square.

So, you can see that the diagonal elements are all real and with a non-negative and all the non-diagonal elements, they satisfy this symmetry relation, conjugate symmetry relation rather that is R_{ij} . That is the i th row and j th column, that entry is R_{ij} and the j th row and i th column that is R_{ji} . So, R_{ij} and R_{ji} , they are conjugate of each other which mean R is a hermitian matrix that you can see in other ways also.

(Refer Slide Time: 12:46)

$$\begin{aligned} \underline{R}^H &= \left[E[x x^H] \right]^H = E[(x x^H)^H] \\ &= E[(x^H)^H x] = E[x x^H] = \underline{R} \end{aligned}$$
$$\underline{R} = E \left[\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} [x_1^* \ x_2^* \ \dots \ x_n^*] \right]$$
$$= \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \dots & R_{nn} \end{bmatrix}$$

What is R^H ? It is E of H . You can put the hermitian operation inside to as you after taking I mean whether you take the expectation of this matrix, and then take their complex conjugate transpose, or you first take the conjugate transpose, and then take the expected value, either you will get the same values, right. So, this is nothing but E of and you have seen a product AB hermitian is nothing but B hermitian. This is $B \times$ hermitian is $B \times$ is A . So, AB hermitian is nothing but B hermitian, that is hermitian of x hermitian, and then this x hermitian and two hermitian gets canceled. You get back x .

So, you get x which is same as R which means R is a hermitian matrix. Hermitian matrices have some very interesting properties. I am coming to that. If your properties which are extensively used in signal processing applications, I will be coming to that. Another will be correlation matrix.

(Refer Slide Time: 14:26)

Handwritten derivation of the covariance matrix formula:

$$\begin{aligned}
 \text{CO-VARIANCE MATRIX} \\
 \underline{C} &= \underline{R} - \underline{M} \\
 &= \underline{R} - \underline{\mu} \underline{\mu}^H \\
 &= E[(\underline{x} \underline{x}^H) - \underline{\mu} \underline{\mu}^H] \\
 &= E\left[\begin{matrix} x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^* & x_2^* & \dots & x_n^* \end{matrix} \right] - \begin{matrix} \mu_1 & \mu_2 & \dots & \mu_n \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^* & \mu_2^* & \dots & \mu_n^* \end{matrix}
 \end{aligned}$$

The element C_{ij} is defined as $E[(x_i - \mu_i)(x_j - \mu_j)^*]$. The final result is $C_{ij} = R_{ij} - \mu_i \mu_j^*$.

You can also define covariance matrix C is nothing but R minus M , where M is a matrix which is nothing but $\mu \mu^H$. That means, we have got $x x^H$ hermitian here. We can also bring $\mu \mu^H$ inside the expectation operation because after all these are constants. So, expected value of this constant will be the constant themselves. So, you can as well bring them inside this. So, now what happens is you have one matrix here, another matrix here, this matrix minus matrix $x x^H$ minus $\mu \mu^H$. You can consider the total matrix.

What is the total matrix? It is a matrix whose say i th row and j th column. What will be the element? It will be $x_i x_j^* - \mu_i \mu_j^*$. What is this? This is a covariance between the two random variables x_i and x_j . After what is covariance, we have seen what C_{ij} is. C_{ij} is nothing but you expand it. So, you get first term $x_i x_j^*$ expected value of which is nothing but R_{ij} . Then, minus μ_i expected value of x_j^* which is μ_j^* . In fact, we have done this already. So, that is why I am not getting into details. $\mu_i \mu_j^* - E[x_i] \mu_j^*$ which is again $\mu_i \mu_j^*$, and the cross term $\mu_i \mu_j^* - E[x_i] \mu_j^*$ which is again $\mu_i \mu_j^*$, and the cross term $\mu_i \mu_j^* - E[x_i] \mu_j^*$ which is again $\mu_i \mu_j^*$. It does not make any difference because these are constants plus $\mu_i \mu_j^*$.

So, this cancels and you will get this, and that is what you have here expected value of $x_i x_j^*$ which is R_{ij} minus $\mu_i \mu_j^*$. So, you get covariance and obviously, covariance I mean if all the elements x_1, x_2 up to x_n , they are uncorrelated with each

other, then we know their covariance. Mutual covariance should be 0. What happens to this covariance matrix? Mutual covariant is C_{ij} equal to 0. Whenever i is not equal to j and whenever i equals j , then it is of course non-zero, because it becomes just a variance that is E of x_i minus μ_i times again x_i minus μ_i star x_i star minus μ_i star, so that it becomes actually E of modulus x_i minus μ_i square which is the variance, right.

(Refer Slide Time: 18:33)

Handwritten notes on a whiteboard:

If x_1, x_2, \dots, x_n : mutually uncorrelated

then $C_{ij} = 0$, if $i \neq j$

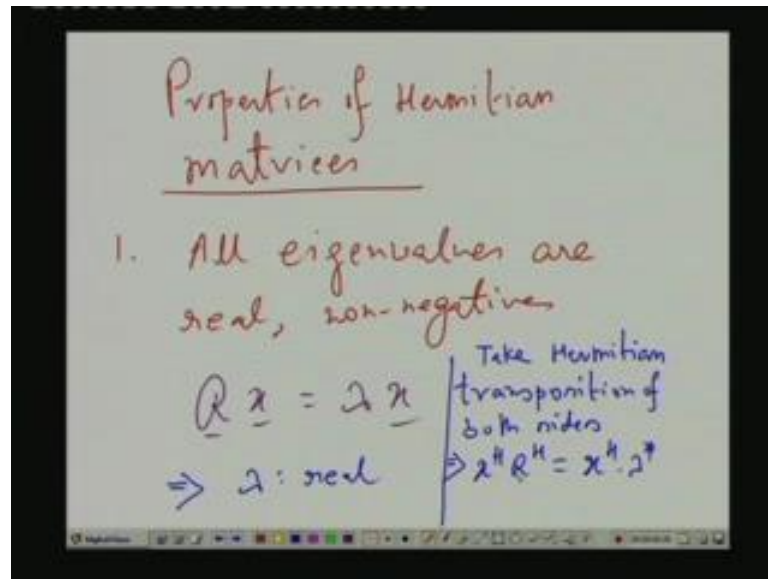
$= \sigma_i^2 = E[(x_i - \mu_i)^2]$, if $i = j$

$C = R - M = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \sigma_n^2 \end{pmatrix}$

In that case, I write it mutually uncorrelated, then C_{ij} equal to 0. If i not equal to j and is equal to what σ_i square which is nothing but E of mod x_i minus μ_i square if i is equal to j . So, that means, C matrix which is R minus M , this is then we take this structure, all non-diagonal entries are 0 that is whenever i is not equal to j , this is 0 and diagonal entries consist of the variances. So, it becomes just a diagonal matrix like this big 0. That means, the upper half and lower half are zeros only diagonal entries which are positive numbers not negative numbers.

So, that defines the covariance matrix for set mutually uncorrelated random variables. In fact, if the means also are 0, if this random variables x_1 to x_n are given to be not only mutually uncorrelated, both are of 0 mean, then of course M becomes all zero matrix except for the diagonal entries. No, it becomes 0, all 0 matrix, because of all the elements are 0. So, in that case, R is simply this.

(Refer Slide Time: 20:38)



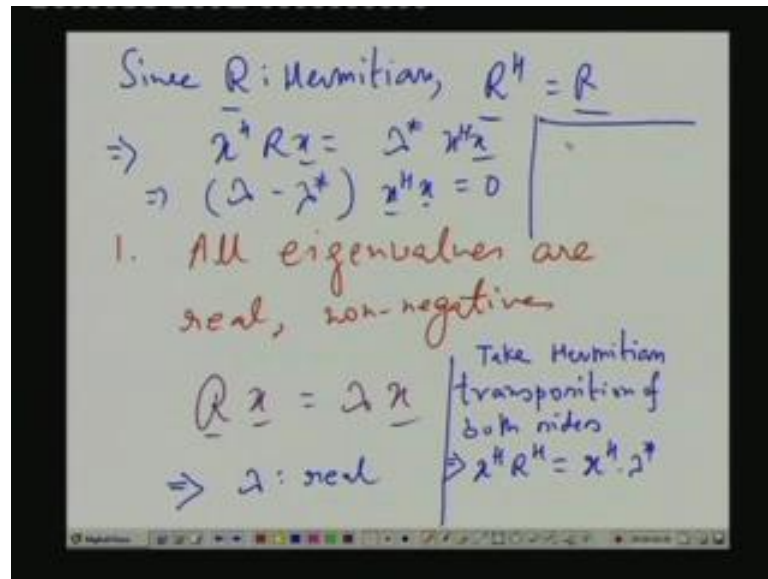
Now, hermitian matrices satisfy some properties. Number one, all Eigen values are real, first real and non-negative. How to show that? That means, suppose $Rx = \lambda x$ is an Eigen vector is λx . Mind you by definition of the Eigen vector, x is a non-zero vector after all. If x is a 0 vector, all 0 vectors are Eigen vectors. Why? It is because R times any matrix times 0 Eigen vector is always 0. Vector is always a zero vector. So, zero vectors is excluded from the definition of Eigen vector. So, that means, whenever I have an equation like this. I say x is a Eigen vector corresponded to the Eigen value λ for the matrix R , and x is there by it is implied that x is a non-zero vector, all right.

We have to show that since R and R hermitian, they are same that R is a hermitian matrix, λ i mean λ is real. It is very easy. You take the hermitian transposition of both sides. It leads to what as you told as I told product of two matrices will take the hermitian transpose. You get $x^H R^H = x^H \lambda^*$ and again λ . You can view λ to be a vector with single entry λ . So, it conjugates. Its hermitian transposition is what you further take the transpose which is λ itself because it is a scalar, and then conjugate.

So, get λ^* and x^H hermitian means you have to take the hermitian of x first, x^H hermitian λ^* , but here it does not matter, because λ^* also is a scalar number. The scalar number you can write to the right of x^H or to the left of x^H , it does not

matter because it is a scalar number, but we have known that we are given rather than R hermitian matrix which means R hermitian transposition is same as R .

(Refer Slide Time: 24:07)



That is since R is hermitian, R^H is same as R which means λ , sorry which means here $x^H R$ is same as $\lambda^* x^H$ and so is a row after all. x^H is a row vector followed by a matrix multiply with a row vector. Similarly, on this side also row vector given by x^H each entry is multiplied by λ^* . So, if I now post multiply this row vector by the column vector say x , here also I bring x , but Rx is we know λx . That means, Rx is λx . So, x^H and λ can be taken outside because it is a scalar.

So, you take these two take both the sides to the same side. So, $\lambda - \lambda^*$ $x^H x$ is 2. Rx is λx Rx is λx . So, you get this put λx here take λ outside. So, $x^H x$ here, also $x^H x$ take that common $\lambda - \lambda^*$. Now, $x^H x$ is what? What is $x^H x$?

(Refer Slide Time: 26:24)

Since R is Hermitian, $R^H = R$

$$\Rightarrow \lambda^* R x = \lambda^* x^H x \Rightarrow (\lambda - \lambda^*) x^H x = 0 \Rightarrow \lambda = \lambda^*$$

1. All eigenvalues are real, non-negative

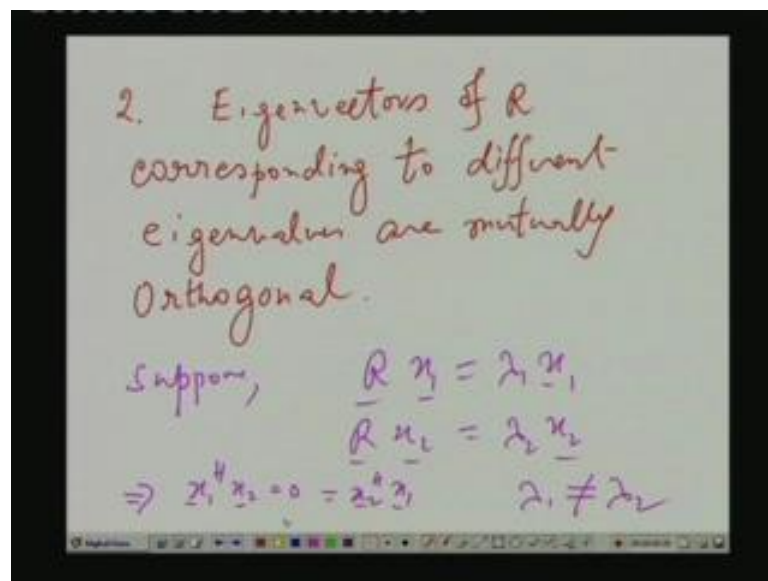
$$x^H x = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \geq 0$$

What is $x^H x$? That is simply if you just I mean expand it, $x^H x$ is a row vector with which entry with each entry conjugated, and then followed by a column vector x carry out the row vector times column vector multiplication you will get which is a real number because you are taking mod, and then square. So, it is always non negative. In fact, it is most of the times positive. If it is 0, each of the terms has to be 0 because you are taking mod of square which means x_1 must be 0, x_2 must be 0 and x_n must be 0. They also having value 0, but point is by definition x cannot be 0 vector. x is an Eigen vector.

If $x^H x$ hermitian see in this product $\lambda - \lambda^*$ times $x^H x$ is 0, $x^H x$ is a scalar $\lambda - \lambda^*$ is a scalar. So, either $\lambda - \lambda^* = 0$ or $x^H x = 0$, but what is $x^H x$. We have seen here mod of x_1 square plus mod of x_2 square plus dot dot dot plus mod of x_n square. So, this is less than greater than equal to 0. It cannot be negative, but if x is given to be an Eigen vector, then it is not only greater equal to 0, it is actually greater than 0 because an Eigen vector means it is a non-zero vector. So, at least one entry has to be non-zero whose mod square is not of course greater than 0. That means, in here $x^H x$ cannot be 0 which means $\lambda = \lambda^*$. So, this leads to $\lambda = \lambda^*$ which means the Eigen values are real.

Next thing you have to show that they are not only real, they are non-negative. We are coming to that. Now, the fact that they are non-negative, I postponed it for the time being. There is some property will be required which we will be coming little later and that property is called positive definite of the correlation matrix. So, I have so far proved that lambda is real. The fact that lambda is also non-negative. It is greater than equal to 0. I will prove it just little later. I will come to the other point property number 2.

(Refer Slide Time: 29:03)



Sorry, it means suppose $R \underline{x}_1 = \lambda_1 \underline{x}_1$, $R \underline{x}_2 = \lambda_2 \underline{x}_2$, and λ_1 and λ_2 , they are not same, that is \underline{x}_1 and \underline{x}_2 , their Eigen vectors, but their Eigen vectors corresponding to two different Eigen values of R . One is λ_1 , another is λ_2 . Then, it says that \underline{x}_1 and \underline{x}_2 are orthogonal meaning if you take this product which is called the inner product $\underline{x}_1^H \underline{x}_2$, this must be equal to 0 which is also true for $\underline{x}_2^H \underline{x}_1$, that is you either make \underline{x}_1 a row vector and then conjugate it, and then multiply the \underline{x}_2 .

So, what are the entries $\underline{x}_1^H \underline{x}_2$ plus? I mean the first entry conjugate times, the first entry of \underline{x}_2 , then first entry conjugate of \underline{x}_1 times, sorry second entry conjugate for \underline{x}_1 times second entry of \underline{x}_2 and like that because sum equal to 0, or alternatively if you can take the hermitian transposition of the left hand side here, you get $\underline{x}_2^H \underline{x}_1$. Obviously, 0 hermitian transpositions give you 0. So, $\underline{x}_1^H \underline{x}_2$ or say $\underline{x}_2^H \underline{x}_1$, they both gives rise

to 0, all right. This is to be proved. In fact, this orthogonal you know transfer a motion of what is called inner product we define for two vectors.

(Refer Slide Time: 31:58)

For any two vectors $\underline{x}, \underline{y}$,
 inner product:
 $\langle \underline{x}, \underline{y} \rangle = \underline{x}^H \underline{y} =$
 $= \sum_{i=1}^p x_i^* y_i \quad [x_1^* \dots x_p^*] \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$
 Suppose,
 $\underline{R} \underline{x}_1 = \lambda_1 \underline{x}_1$
 $\underline{R} \underline{x}_2 = \lambda_2 \underline{x}_2$
 $\Rightarrow \underline{x}_1^H \underline{x}_2 = 0 = \underline{x}_2^H \underline{x}_1 \quad \lambda_1 \neq \lambda_2$

Inner product is like what in three-dimensional vault. You had for dot product between two vectors, then take two vectors and get a real number out of the dot product. Its satisfaction basic properties for two vectors actually define like this. There you write 1 vector, then dot y by, but widely used notation is this inner product or dot product instead of having just \underline{x} dot \underline{y} we write within these two things symbols $\underline{x}, \underline{y}$ which is nothing but \underline{x} hermitian \underline{y} . So, that means, one is this vector \underline{x} 1 star up to say \underline{x} star. If there are p entries, another is \underline{y} 1 dot dot \underline{y} p which is nothing but summation x_i star y_i i equal to 1 to p . If the inner product is 0, then we say \underline{x} and \underline{y} are orthogonal. We have to prove that two Eigen vectors \underline{x}_1 \underline{x}_2 corresponding to different Eigen values λ_1 and λ_2 of \underline{R} , they are mutually orthogonal. Now, how to show that?

(Refer Slide Time: 33:32)

$$\left. \begin{aligned} R \underline{x}_1 &= \lambda_1 \underline{x}_1 \\ R \underline{x}_2 &= \lambda_2 \underline{x}_2 \end{aligned} \right\} \lambda_1 \neq \lambda_2$$

$$\underline{x}_1^H R^H = \lambda_1^* \underline{x}_1^H \Rightarrow \underline{x}_1^H \underline{x}_2 = 0$$

$$\Rightarrow \underline{x}_1^H R = \lambda_1^* \underline{x}_1^H$$

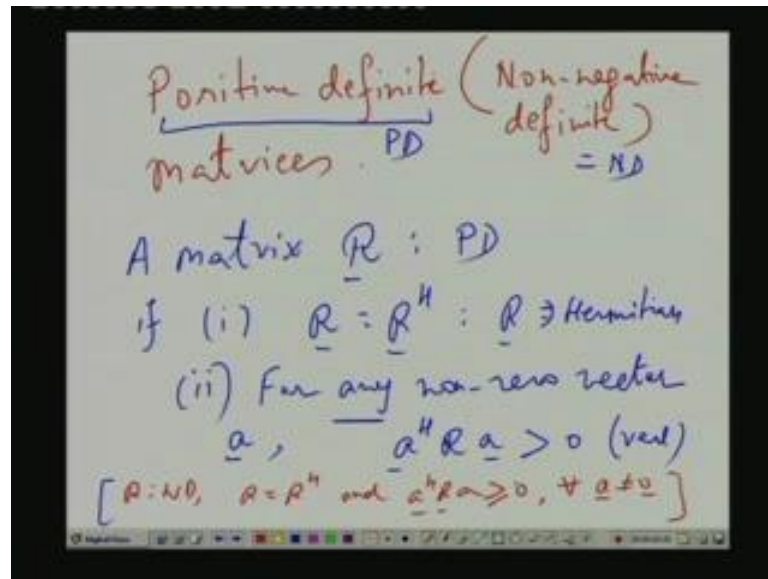
$$\underline{x}_1^H R \underline{x}_2 = \lambda_1^* \underline{x}_1^H \underline{x}_2$$

$$\Rightarrow (\lambda_2 - \lambda_1^*) \underline{x}_1^H \underline{x}_2 = 0$$

As I said $R \times 1$, it is given $\lambda_1 \neq \lambda_2$. Now, you take the hermitian transposition of the first equation that is now you take hermitian transposition here. What you get is $\underline{x}_1^H R$ hermitian, but R hermitian is same as R . Do not forget that is same as \underline{x}_1^H hermitian λ_1 hermitian, but λ_1 is a scalar. Its hermitian is conjugate of λ_1 , and you can always write it in the beginning only because it is a scalar number. So, $\lambda_1^* \underline{x}_1^H$ also means as before, sorry as before R^H and R are same. So, I will just write in like this λ_1^* .

So, I now pre-post multiply this row vector with \underline{x}_2 . So, I bring \underline{x}_2 here. I wrote the same here also, but λ_1 as I told this is very important. λ_1 is Eigen value is real. So, λ_1^* is same as λ_1 . Is it not? It is because they are real. I have just now proved a while back that is $\lambda_1 \lambda_1^*$, they are same real. So, I write $\lambda_1^* \underline{x}_1^H$, and then \underline{x}_2 is present, and $R \underline{x}_2$ is given in the second equation is $\lambda_2 \underline{x}_2$. Replace $R \underline{x}_2$ by $\lambda_2 \underline{x}_2$, take all the terms to one side. So, you get $\lambda_2 \underline{x}_1^H \underline{x}_2 - \lambda_1^* \underline{x}_1^H \underline{x}_2 = 0$, which means since λ_1 and λ_2 are not same, this means only possibility is that this is equal to 0.

(Refer Slide Time: 36:19)

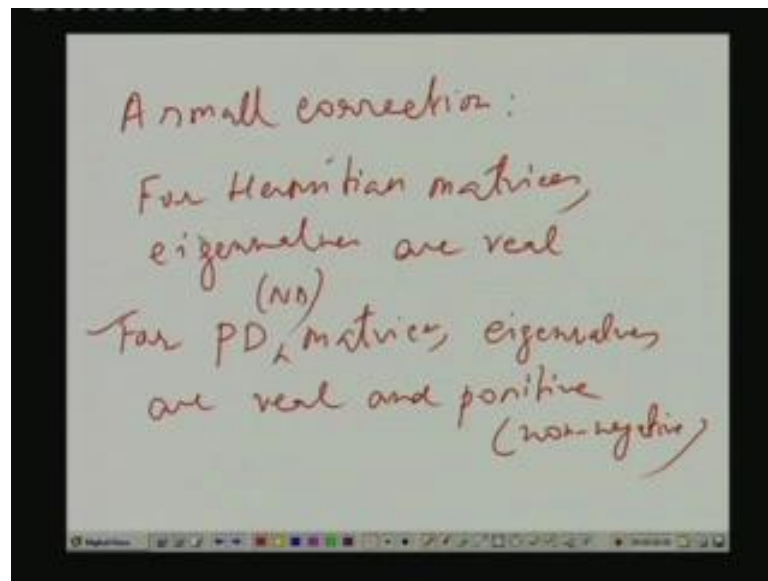


Now, I come to positive definite or non-negative is a more general case non-negative definite matrices. Here, positive definite I call it PD and non-negative definite, I call it ND. So, this is ND abbreviation and this is PDR. So, matrix R is PD. First positive definite if number 1, R equal to RH that is R is hermitian, and secondly for any non zero, any non-zero vector, the scalar number a hermitian Ra is greater than 0. Not only this is of course this is real and greater than 0, and then I add R ND. If R is non-negative definite condition, one remain same that is R has to be R hermitian and a hermitian Ra instead of greater than 0, it will just become greater than equal to 0 for all non-zero a.

So, only difference is this in there is positive. You see in both the cases of positive definite as well as non-negative definite matrices, R is hermitian in the case of positive definite matrices. So, what we have is that if you take any non-zero vector a, then a hermitian Ra is a scalar number which is greater than 0 of course, and if it is non-negative definite matrix, then a hermitian Ra is greater than equal to 0 still real. So, I will consider non-negative definite matrices, and then in fact you know because this is more general, I mean you can easily consider positive definite case as an extension of that. Then, my statement is R correlation matrix that is so long I was considering at general hermitian matrix, there were no connection with correlation matrices and all that. I will consider a general hermitian matrix and got those properties.

Now, I am coming to R or may I first prove that if R is non-negative definite, actually I made a small mistake. I mean I said that for a hermitian matrix, Eigen values are non-negative definite, i mean greater than equal to 0. That is not true as such. Actually at that time, I had in mind what is called correlation matrix which is always non-negative definite that I am coming now. So, I will just correct my statement.

(Refer Slide Time: 40:54)



Small correction actually for hermitian matrices values are real for PD matrices, Eigen values are PD, and then within bracket ND, either PD or ND Eigen values are real and positive within bracket non-negative, that is if it is a non-negative definite matrix, Eigen values are real. Not only that they are greater than equal to 0, that is they are not negative, but if it is positive definite matrix, Eigen values are of course real, and they are strictly positive, but these happens if the matrix is positive definite or non-negative definite. That is not just hermitian, but that additional condition is present. If it is hermitian, we can only say that Eigen values are real.

So, previously I made a statement which was not fully correct because I just considered hermitian matrices. Actually I had in mind what is called correlation matrices. Correlation matrices I will now show that they are actually non-negative definite matrices. Now, how to show first that if a matrix is non-negative definite, Eigen value is

of course they are real. We have shown, but they are also non-negative that is greater than equal to 0.

(Refer Slide Time: 42:51)

The image shows a whiteboard with the following handwritten text:

$$R: \text{non-negative definite}$$

$$R \underline{x} = \lambda \underline{x}$$

$$0 \leq \underline{x}^H R \underline{x} = \lambda \underline{x}^H \underline{x}$$

$$\Rightarrow \lambda \underline{x}^H \underline{x} \geq 0 \Rightarrow \lambda \geq 0$$

Now, suppose R is given to be non-negative definite, now Rx as you have seen is λx suppose, then I just pre-multiply it by x^H , and then what I get here is $x^H R x$. Now, by definition x is an Eigen vector of R which means x is a non-zero vector, and it is R is given to be non-negative definite any $x^H R x$ for a non-zero x must be greater than equal to 0 must be, it is non-negative definite. If it is strictly positive definite, it is strictly greater than 0. If it is non-negative definite, it is greater than equal to 0 which means $\lambda x^H x$ is greater than equal to 0 $\lambda x^H x$, but we have also seen what is $x^H x$, that is you take each component of x to the mod of that and square and sum.

So, that is always greater than 0 because x is non-zero vector. So, at least one entry will be non-zero. So, its mod value is positive, square is positive which means $x^H x$ cannot be 0 which means it cannot be negative or cannot be 0 which means λ has to be greater than equal to 0 say $x^H x$ is always positive because x is Eigen vector. So, not that x can be 0 that there will be at least one entry of x which is non-zero. So, mod square of that is positive which means $x^H x$ is always positive, strictly positive. So, λ times that if it is greater than equal to 0, the only competition is this. So, if this is greater than equal to 0, this only means λ has to be greater than equal

to 0 and lambda is real. You have already seen, and if it is positive definite that instead of greater than equal to, we have that symbol greater than. So, in that case lambda is greater than 0.

(Refer Slide Time: 45:14)

$R = \text{correlation matrix}$
 R is Non-negative Definite
 Take $\underline{a} \neq 0$
 $\underline{a}^H R \underline{a} = E[(\underline{a}^H \underline{z})^2]$
 $= \underline{a}^H E[\underline{z} \underline{z}^H] \underline{a} \geq 0$
 $= E[(\underline{a}^H \underline{x})(\underline{x}^H \underline{a})]$
 $= E[(\underline{a}^H \underline{z})(\underline{a}^H \underline{z})^H]$

Now, I am coming to my main issues which promote me actually to get into the properties of hermitian matrices or positive definite matrices and all that, that is R which is a correlation matrix is non-negative definite. First part of this is that R must be hermitian which you have already proved. We have dealt with that link. So, I am not coming to that. You have come to the second property that is for any non-zero vector x or a hermitian Ra is a number which is non-negative, either zero or positive, but it is a non-negative number.

How to show that? So, you take a, to be a non-zero vector take as before a, is a non-zero vector. That is a vector whose at least one entry is a non-zero element. It is not that all entries are 0. Then, what is a hermitian Ra ? For any such non-zero a, if a hermitian Ra is first real greater than equal to 0, then we can say that R is non-negative definite because R is hermitian. We have already proved that. Now, what is a hermitian a? You can write, but mind you a is constant a is not random. You have to just pick up some vector a by your choice, but there is no random element in it. Now, R is as we now E of, and then a. Since, the elements of a are constant, you can bring them within the expectation

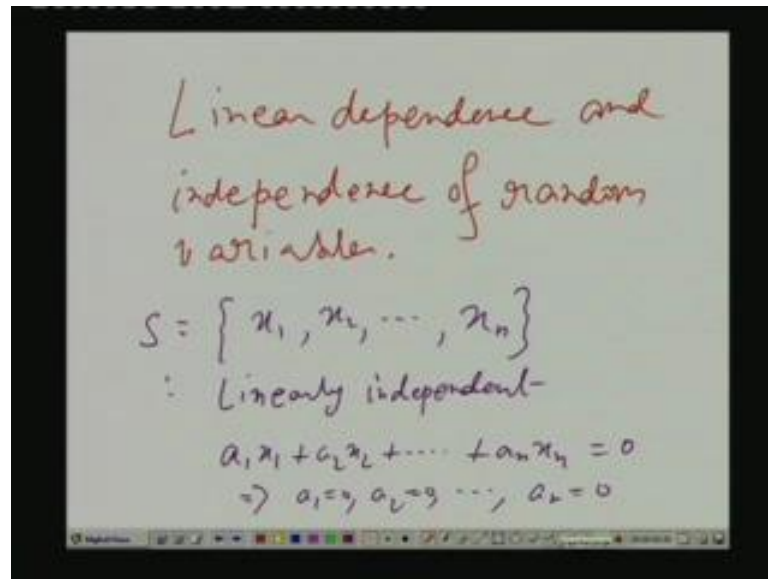
operations. It does not change anything. You can write it like this a hermitian H may be you can put it in the bracket, and then x hermitian H .

Now, a hermitian x this is a scalar number, right row vector column vector and you see what is x hermitian a . This is nothing but complex conjugate or this is a scalar number and if you take the hermitian transposition, just a minute this is what you can write. X hermitian a is nothing but hermitian transposition of a hermitian x . If really a hermitian x and take the hermitian transposition of that, x hermitian comes first a hermitian hermitian is that comes second and remember that a hermitian x is a scalar number. So, it is hermitian transposition is nothing but there is no transposition involved here. Only complex conjugation which means this gives rise to E of mod a Hx square which means this is a real number first because this is mod involved and mod square.

So, it cannot be expected values of this cannot be negative. It is greater than equal to 0. So, it is non-negative definite. If it so happens, we can ensure that it is always greater than 0. Then, we can also say R is positive definite. Now, when can we do that? Before that consider this if I say that expected value of a random variable say x or say z expected value of z is 0, that means, z is a random variable which always takes the values 0, and then only its expected value will be 0.

Now, if R is non-negative definite and not positive definite as such, that means expected value of this mod square of aHx , this is as such we know this has to be greater than equal to zero, but if it is non-negative definite, it means that you can find at least one a one non-zero a for which aHx is random variable. It takes zero value always, then only its mod square expected will be 0. I repeat again if it is R is given to be strictly non-negative definite and not just positive definite. That means, for some a that means, it is not just greater than 0. It is greater than equal to 0 which means sometimes it can be equal to 0 also, but when it equals to zero, that means for some a that is there you can find at least 1 a if not more for which a is non-zero for which aHx answer to be zero always. Then only its mod square upon expectation also would always gives rise to 0, and you can put an equal to 0.

(Refer Slide Time: 50:18)



When does that happen? That leads to linear dependence and independence is set x_1 , sorry is linearly independent. If it is linearly independent, that means, there is no linear relation involved in these random variables means any of the random variable here of x_1 to x_n can be written as a linear combination of the rest. Then, it is called a linearly independent relation which means actually using (()) theory or linear (()), it appears that you really want to form a combination like a 1×1 plus a 2×3 plus dot dot dot and x_n equal to 0 is only way you can do is that a 1 equal to 0, a 2 equal to 0 dot dot dot an equal to 0. That is you can easily see.

Suppose what I mean here is this if you really form an equation like equal to 0. That means if you linearly combine these linear variables and equate to a zero random variable, right hand side is a zero random variable which means it always takes the value 0. If you really want to find out such a 1 , such a 2 and such an, which maintains its equality, then if they are linearly independent, the only way we can do is by putting a 1 equal to 0, a 2 is 0 and an equal to 0. According to b, under this condition you can never express any of the random variables say x_1 as a linear combination of the rest.

Suppose just to give an example, suppose a 1 and a 2 , they are non-zero and rest are 0, that means a 1 times x_1 plus a 2 times x_2 will be equal to 0, say x_1 will be you can always write x_1 in terms of x_2 , or say a 1 not equal to 0, a 2 not equal to 0, a 3 not equal to 0 and rest are 0. You can then write a 1×1 plus a 2×2 plus a 3×3 equal to 0, other

elements are already become 0. So, x_1 then you can keep on the left hand side, right hand side I may put x_2 and x_3 terms. So, you can write x_1 in terms of x_2 and x_3 as a linear relation so on and so forth. In fact, suppose if a_1 is only one element, a_1 is non-zero, other side zero in that case you have got just a 1×1 is 0. That means since a_1 is non-zero, x_1 equal to 0, that is x_1 is the random variable, and a zero random variable you know if you have a zero random variable in this set S, then the set is not linearly independent because zero random variable can always be written as linear combination of the rest. Just multiply each of the random variable by this scalar zero, zero times x_2 plus zero times x_3 plus dot dot dot zero times x_n . That makes it a zero random variable.

So, just to test whether a set of linear, I mean random variables are linearly independent or not, form an equation like this with a_1, a_2, \dots, a_n being scalars equated to Zero random variables. See what choice of a_1 to a_n can give rise to can maintain this equality, and if it is such that only possibility is by taking a_1 equal to 0, a_2 equal to after a_n equal to 0, then you have this. You can say they are linearly independent. If they are linearly independent, then you can also see that any subset is linearly independent. That is easily seen because in the subset, you can express some element as a linear combination of the rest, then that applies for the entire set also because a linearly independent linearly dependent set. Linearly dependent subset is a bigger set that makes the overall set also linearly dependent. On the other hand, if you can find out some coefficients a_1 to a_n when not all are 0, but this equality is maintained, then you say this set is linearly dependent set which means some elements i at least one element can be expressed as a linear combination of either the whole of the rest or part of the rest. Now, in the term we are running short of time.

(Refer Slide Time: 56:00)

$$S = \{x_1, \dots, x_n\} : \text{lin. ind.}$$
$$E [|a_1 x_1 + \dots + a_n x_n|^2] > 0$$

if $a_1 \neq 0$
 $a_2 \neq 0 \dots a_n \neq 0$

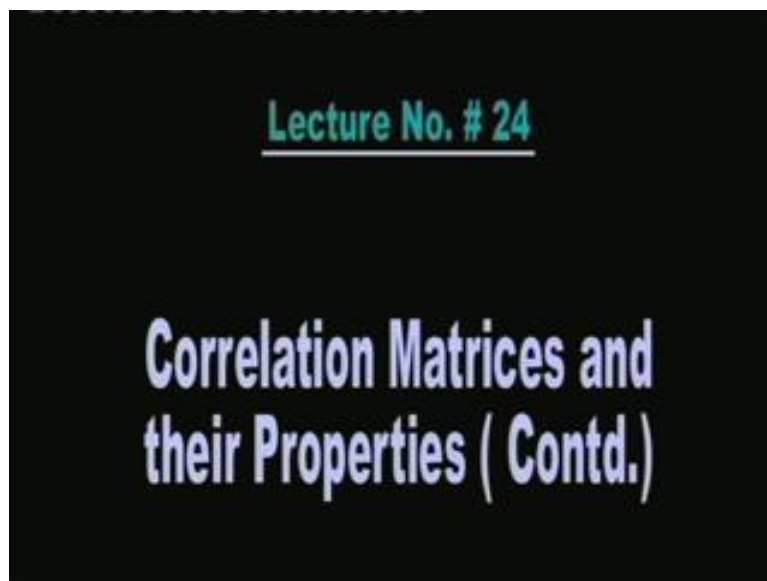
So, I do ending of soon that means linearly independent; that means expected value of any x 1 you take the mod square. This has to be greater than 0. If a_1 not equal to 0, a_2 not equal to 0 dot dot dot a_n not equal to 0 because if they are not equal to 0, this cannot be a random variable and if it is not zero random variable, its expected value of the mod square has to be greater than 0. So, I stop here today. From here I will continue in the next class. So, I will continue these remaining properties of the correlation matrices and all that.

So, thank you very much.

(Refer Slide Time: 50:57)



(Refer Slide Time: 57:06)



So, in the last class, we were discussing these correlation matrices. We discussed some properties of hermitian matrices which are also I mean positive definite matrices. We will find out hermitian as such not even positive definite, they are always real, and their Eigen vectors are orthogonal to each other. We have found out. Then, if in addition the matrix is given to be positive definite, then Eigen values are also positive. If it is non-negative definite, they are non-negative. This we have seen.

Then, we are discussing the linear dependence and independence of random variables and all that. I just start from there. I continue with the previous discussion on the correlation matrices first. Do you remember our previous treatment on what is called vector space of random variables, where each random variable was treated as a vector in an abstract vector space of all possible random variables in the world, and there we defined addition of two random variables, such addition of two vectors. What is meant by zero vectors? That is random variable which always takes zero values.

(Refer Slide Time: 58:32)

$$\begin{aligned}
 T^H T &= \begin{bmatrix} e_1^H \\ e_2^H \\ \vdots \\ e_n^H \end{bmatrix} \begin{bmatrix} e_1 & \dots & e_n \\ \vdots & & \vdots \end{bmatrix} \\
 &= I \quad \Rightarrow \quad T^H = T^{-1} \\
 R &= T D T^H \quad T T^H = I \\
 \det[R] &= \det[T] \cdot \det[D] \cdot \det[T^{-1}] \\
 &= \lambda_1 \lambda_2 \dots \lambda_n
 \end{aligned}$$

Then negative of a random variable means secondary random variable, which takes negative of the original random variable and likewise is just in that frame work we can discuss. Two entries of T times determinant of D times determinant of TH which is nothing but T inverse, and these two cancels. Determinant of T, determinant of T inverse is cancelled. So, you get back determinant of D. What is determinant of D? It is the product of Eigen values $\lambda_1 \lambda_2 \dots \lambda_n$. As you have seen if the Eigen values are first real, so determinant of R is always real. That is done. I mean if they are all positive, if the Eigen matrix is positive definite, then each is positive. So, determinant is positive. So, that is all for today. We continue from here in the next class.

Thank you very much.