Introduction To Adaptive Signal Processing

Prof. Mrityunjoy Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture No # 04
Statistical Impedance, Covariance Matrices

So, we start from where we ended last time. Suppose, suppose there are two jointly random variables ok. So, this part is x and this is dx, this is y and this is dy. We have seen this thing the meaning of this we know y by x. It means when we constraint capital X to lie from here to here under that condition, under that constraint, what is the probability of capital Y lying from here to here? So, whether here or here or here, does not matter that probability. Now, suppose capital X and Y are such there is absolutely no relation between them, as I told you, one could be temperature capital X maybe another one is, another one is say maybe rupee value of dollar, So, absolutely no relation.

So, under such cases if physically Y does not depend on X at all. So, it does not matter whether I constraint capital X to lie, to lie from small x to $x + dx$ or maybe x' to again dx whether here ok. Because they are so there are no relation between them. So, it does not matter whether temperature is from here to here or here to here ok.

This will always be same in both the locations as long as dy is fixed ok. So, it will be like the chance of capital Y falling from small y to $y + dy$. It will be that it will be independent of capital X. Because physically if they are such that there is no relation between them, So, there is no point, there is no meaning of the statement that I am constraining capital X to lie from here to here and then measuring Y and capital Y chances of that you know falling from small y to $y + dy$ is this one.

$$p_{Y/X}(y/x)dy$$

So, that will be independent of where capital X is, So, this will then turn out to be, this is independent of that, it will simply be this.

Just general as we have seen earlier just probability of capital Y lying from small y to $y + dy$ that was this. The conditioning will have no effect. In that case p x y as we have seen earlier, this you could write as sorry free variable, this you have seen already, the constraining variable capital X, so, now, it will be the probability density of this. So, this is as you have seen under the above assumption that there is no you know dependence of capital Y or X and vice versa. This will be simply p y p x x.

$$P_{X,Y}(x,y) = P_{Y/X}(y/x)dy$$

$$= P_Y(y)P_X(x)$$

There is a product of individual probability densities. Individual probability densities also called marginal densities and in this case we say X and Y statistically independent. I abbreviate it as S I statistically independent. Then the joint density is a product of marginal densities or this conditional density has no effect of the condition, So, if the conditioning variable capital X that goes. It becomes just the marginal density of the free variable Y all right.

You can generalize it to case like you know three variables. Suppose now you have got three random variables X, this is Y, this is Z and x, this much is dx, y, this much is dy and z, this much is dz. In that case also. My drawing is not good. So, let me erase this.

This much is $d\ x$, this is y and then sorry. This is $d\ y$. Now, I got it. So, here suppose I consider this thing $p\ z$ is the free variable, suppose at $x\ y$, as an example only, $d\ z$. We have seen it all what does it mean, if I constraint capital X and Y to lie, capital X lie from here to here, capital Y line from here to here.

That is this plane whether here or here or here. Subject to that the probability of capital Z

lying from small z height to height small $z + dz$. That was equal to this proportional to $d\ z$ and proportional to constant. But suppose capital Z is such a random variable it has nothing to do with $x\ y$, no relation. So, it will not depend on the constraint.

$$P_{Z/XY}(z/xy)dz$$

$$P_Z(z)dz$$

It will be simply equal to just p z z d z, marginal density. In that case we will say, sorry, in that case we will say, z is SI statistical independence with $x\ y$ and in this case we will have p x y z, you can write as p z by x y times p density of, joint density of x y and this is nothing but p z only now all right.

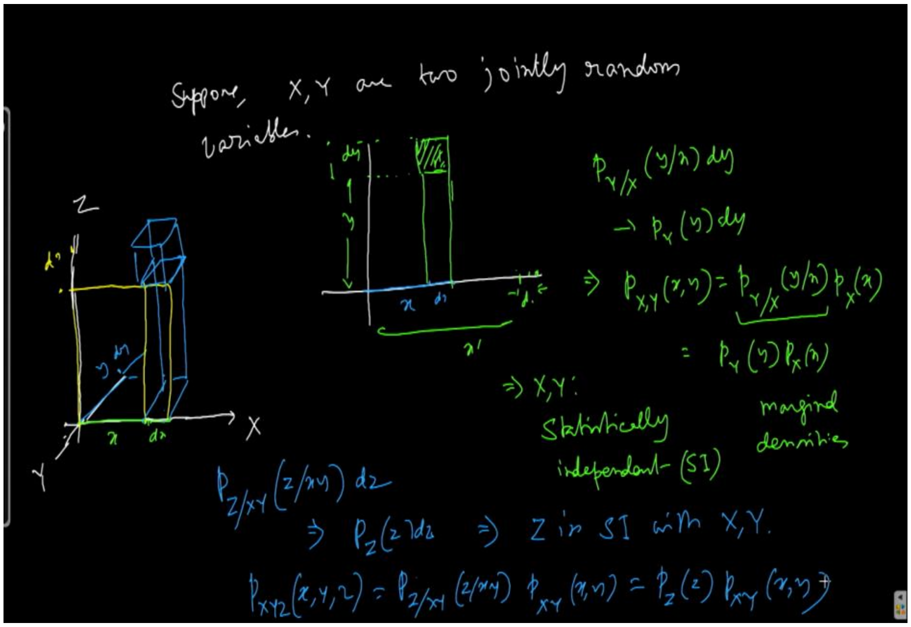
$$P_{XYZ}(x, y, z) = P_{Z/XY}(z/xy)P_{XY}(x, y)$$

$$= P_Z(z)P_{XY}(x, y)$$

So, you can write like this. That means, if they are, if out of the all the random variables, some of them are statistically independent of the rest. So, you take the joint density or in this case marginal density because you have already one, but you take the joint density of one set here and multiplied by the joint density of the other set here.

The two sets are statistically independent. If further capital X also and Y also are statistically independent between themselves then this again further you can break as p capital X of small x into p capital Y is small y. So, the three are statistically independent in that case. So, depending on the case you can break up like this all right. This is one thing.

(Refer Slide Time: 11:47)



Next, now we consider one expression that was considered earlier, covariance between two random variables X and Y. Covariance - it was denoted by c x y, expected value of, minus the mean y, minus the mean, its physical meaning was explained that time, this is the ac part, incremental part around the mean this is incremental part around the mean. If x and y are highly related to each other, either both will go up above their mean together or below their mean together. So, product will be positive in this case, plus into plus- plus, minus into minus till plus and only average out, you get a good number positive number. Or it could be such when it goes out, maybe it goes down and vice versa.

$$C_{X,Y} = E_{XY}[(X-\mu_X)(Y-\mu_Y)]$$

So, in this case if you average out you will get a good negative number of good magnitude. But if there is no relation between them then sometimes this increment could be positive and some then it is positive and again when this is positive sometimes it could be negative and vice versa. So, when you take all cases and product and average you know it may cancel each other and you get close to 0. So, that you give us an idea about the degree of correlation between x and y, if covariance is high they are highly correlated, high means high magnitude. If covariance is 0 they are not correlated they are called uncorrelated all right.

$$C_{X,Y} = E_{XY}[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y]$$

So, now, this E is actually there are two random variables. So, now, I can put $x\ y$ and E is a linear operator. So, if I now break it, x into y minus mu x y minus mu y y plus mu x mu y. So, it will be E x y on this, then minus E x y on this, then minus E x y on this, plus E x y on this. So, if you apply E x y on this, so, product of the two random variable, they are average expected value.

It is called correlation. It is denoted by correlation. And then E x y on this, of which mu x is a constant, E x y means what multiplying by the joint density and integrate. So, mu x is a constant not random it will go out. So, it will be mu x times E x y of y.

But I told you I in previous lecture, I have shown that if there is a function and the function here is just y itself, it is a function of one variable and you have got more than that variable here and you have got other variables also with respect to which you carry out the expectation. Then those variables simply drop, So, x will drop. So, this is equivalent to E y y which is equal to mean of y that is mu y. Similarly, E x y of, 1 minute, x mu y, it is just a minute, it is x into mu x into y mu y into x this will be x. So, again mu y will come out.

E x y we work on x. x is the function, but it has only one variable x. So, y will drop out. It will be E x x which is mu x. So, mu x and then E x y on this. So, this times the joint density integrated.

There is a meaning of E x y of something, is a constant. E x y that means, this into the joint density p of capital X comma y bracket small x small y dx dy integral. So, this will come out of the integral because this is constant. And that integral will be 1, because the double integral of that joint probability density. So, it will be just mu x mu y.

$$C_{X,Y} = E_{XY}[XY] - \mu_X E_{XY}[Y] - \mu_Y \mu_X + \mu_X \mu_Y$$

$$\equiv E_Y[Y]$$

$$= \mu_Y$$

So, minus mu x mu y, minus mu x mu y, plus mu x mu y. So, it will give rise to R x y minus mu x mu y.

$$C_{X,Y} = r_{XY} - \mu_X \mu_Y$$

This means if x y are uncorrelated then $C_{XY} = 0$, which means, it could be 0, correlation is product of the means. And often we deal with random variables which have 0 mean. So, if mean, one of the means is 0 or both the means are 0, then uncorrelated, which means this that will also mean, correlation also 0, covariance 0, correlation also 0 all right.

But in general, uncorrelated means covariance 0 and correlation is a product of the means. Only if the means are, at least one mean is 0 then correlation also is 0, But not always, this is the meaning of uncorrelated.

$$r_{XY} = \mu_X \mu_Y$$

Now suppose x y are statistically independent that means, this joint density is product of the marginal densities. It was taught in the previous class. In that case, r_{XY} which is $r_{XY}$is a expected value of the product, so, double integral, small x, sorry this is capital So, capital X takes values small x, capital Y takes value small y into the joint density dx dy.

$$P_{XY}(x, y) = P_X(x) P_Y(y)$$

$$r_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \, P_{XY}(x, y) dx \, dy$$

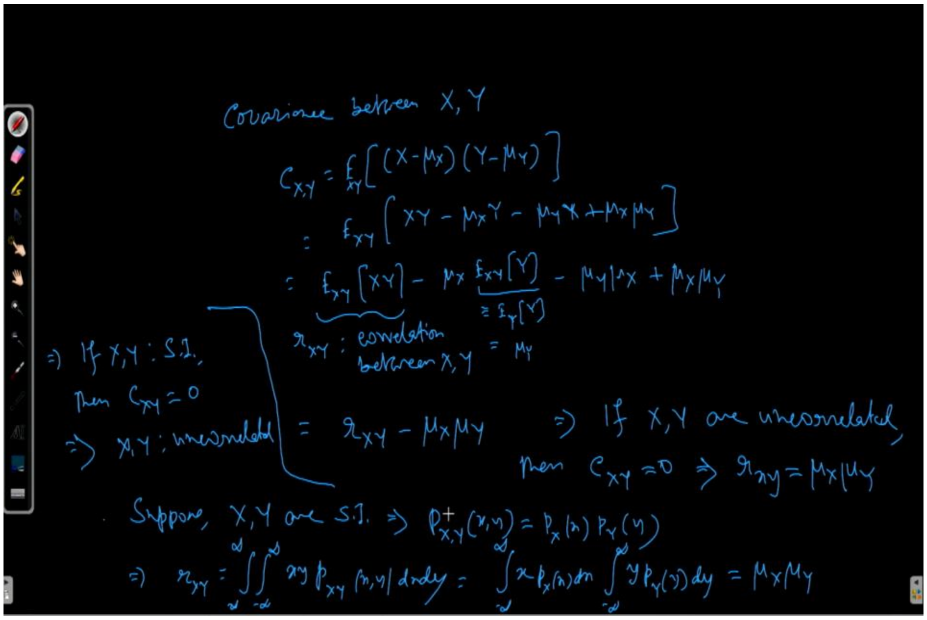
So, there is a general formula, this is a function of two variables, which is the product itself multiplied by the joint density double integral. So, the joint density if you can break like this because they are statistically independent then we have one integral, in which x will occur at this p and dx, another integral, but this is the mean expected value of x, this is the expected value of y. So, this is mu x mu y.

$$= \int_{-\infty}^{\infty} xP_X(x)dx \int_{-\infty}^{\infty} yP_Y(y)dy = \mu_x\mu_y$$

So, if they are statistically independent r_{XY} is the correlation of the product of the means means they are uncorrelated the covariance is 0. This means if x y are statistically independent then $C_{XY} = 0$ because r_{XY} is product of the mean as required here means x y uncorrelated.

So, if they are statistically independent that is if the joint density can be broken, it can be as a product of the individual marginal densities probability densities then the covariance is 0, correlation is the product of the means and they are uncorrelated. But the converse is not always true that is suppose this is given that covariance is 0 or equivalently r_{XY} correlation is product of the means, that does not necessarily mean that the joint density you can write as a product of the individual densities. That happens only once when capital X and Y are called jointly Gaussian random variables, that I will consider later, but not in general. This always always true if statistically independent they are uncorrelated, but if they are uncorrelated it is not guaranteed that they will be statistically independent, that is the joint density can be written as a product of marginal densities ok. That depends on that particular density and there is a particular probability density form called Gaussian Gauss Gaussian.

(Refer Slide Time: 21:46)



So, joint Gaussian for two random variables that we will see later ok. Now with all this background I consider multiple random variables not just X or Y or X Y or X Y Z I will have seen general maybe p number of or n number of random variables joint random variables. So, let X 1 X 2 instead of Y, X 1 X 2 dot dot X maybe p, p joint real continuous random variables that is in every experiment you are observing these separately. One could be temperature, one could be humidity, one could be other things, you know, I mean all together, there may be some relation between they may not be, they are observing all right. So, they have a joint density or all those things.

$$X_1, X_2, \ldots X_P$$

Everybody has a mean, so, E X 1, E X I of X I is a mean of that mu I. So, X 1 has been mu 1, X 2 has been mu 2 dot dot dot all right. and I form a vector small x bar, underscore, I put these variables in a vector form, in a stack all right. Then from each and also I put the means in a vector form, then consider this X minus mu, that is X 1 random variable minus mu 1, X 2 minus mu 2, dot dot dot. So, increment or the incremental part around the mean, it can be positive negative like the AC component this is the DC this is the, mu is the DC average and this is the AC around the mean.

$$E_{X_i}(x_i) = \mu_i$$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_p \end{bmatrix}$$

$$\underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \mu_p \end{bmatrix}$$

$$\underline{x} - \underline{\mu} = \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \cdot \\ \cdot \\ X_P - \mu_p \end{bmatrix}$$

So, fluctuating sometimes positive sometimes negative same here same here ok. Now if I consider this thing X minus mu vector as it is, this column vector, into its transpose, it will be a matrix, column into row. So, ith row, any ith row and jth column, maybe jth column, this limit will be what? ith So, Ith guy and when it becomes row, jth guy that is X_i, X_j minus μ_j, their product, if you see the way this matrix is form this is a vector it is row. So, first element into first element of this, first element into second element of this, first element into third element of this, that is of the first row formed. Then second element into first element, second element into second element, second element into third element, dot there is a second.

So, ith row is formed how ith element times first guy that is here, first ith element times second guy that is here second like that. So, ith times the jth ok. So, this will be X_i minus μ_i into X_j minus μ_j all right and if I apply now expected value on this, expected value with respect to all of them, $X_1, X_2, \ldots X_P$. So, this E will work here, but you see only two random variables $X_i X_j$ present. So, as I told you all others will drop off only X_i and X_j will remain.

$$E_{X_1,X_2,\ldots X_P}[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})^2]$$

So, it will be E of just $X_i X_j$, all other random variables will go, that I have told already shown earlier. Only those who are present in this function on them those variables only will occur. So, we will get and this is nothing, but covariance between these two random variables. So, $C_{X_iX_j}$ all right.

$$E_{X_iX_j}(X_i - \mu_i)(X_j - \mu_j) = C_{X_iX_j}$$

So, this matrix is called covariance matrix. Matrices are are shown by me by upper case letter with an underline. Vectors column vectors are shown by me by lower case letter with an underline. But if there is no underline then it is a scalar it could be upper case lower case does not matter. So, I repeat again capital letter underscore - matrix, lower case letter underscore - vector and if it is not underscore it can be lower case upper case, but it is your is a scalar fine.

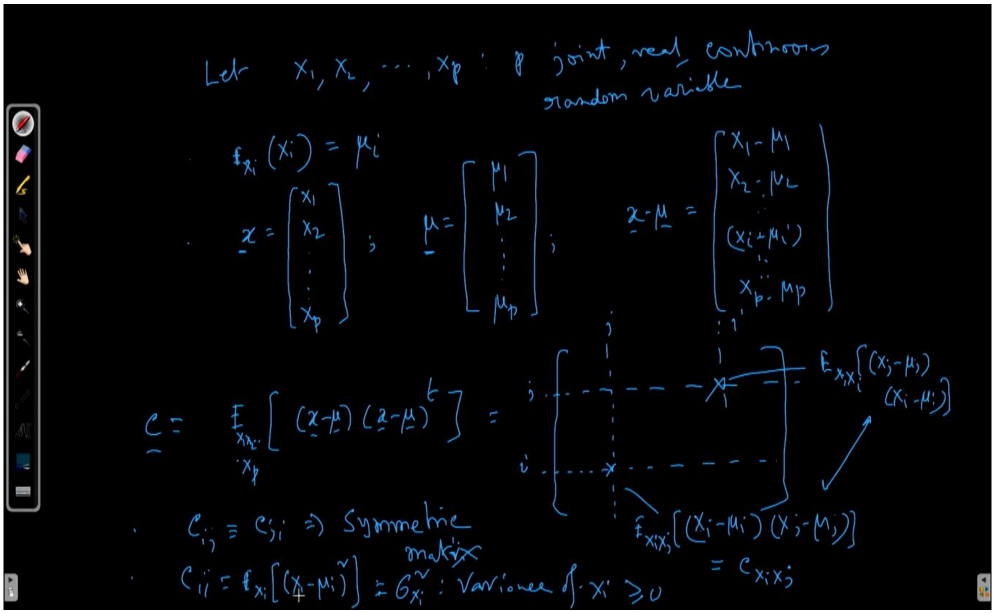
So, this is the covariance matrix. Now you see one thing if I now take the j comma ith row, it was i comma j by the same way j comma ith in the same token, you will have this will be E again only first j comma i first take this product, So, X_j minus μ_j, now that will multiply X_i minus μ_i and only these two random variables will survive here, others will drop off. So, it is this, but these two things are same. X_i minus μ_i, X_i minus μ_i, it is first here, it is second here. X_j minus μ_j coming second it is coming first, but product is same.

So, they are same. So, C is such a matrix C i comma j is same as C j comma i. So, it is a symmetric matrix. Another one C i comma i there is a diagonal ith row ith column that will be both are same, that will be E that is X_i only X_iX_i is X_i and X_i minus μ_i, X_i minus μ_i, j and i are same now. So, X_i minus μ_i whole square which is the variance, sigma X_i square variance of X_i which is greater than equal to 0 variance cannot be negative because this is square and then you are averaging.

$$C_{ij} = C_{ji}$$

$$C_{ii} = E_{X_i}[(X_i - \mu_i)^2] = \sigma_{X_i}{}^2 \geq 0$$

(Refer Slide Time: 30:09)



So, then this cannot be negative ok. So, I stop here we will discuss properties of correlation matrix or covariance matrices in the next class. Thank you very much.