Indian Institute of Technology Madras

NPTEL
NATIONAL PROGRAMME ON TECHNOLOGY ENHANCED LEARNING

Pattern Recognition

Module 06

Lecture 02

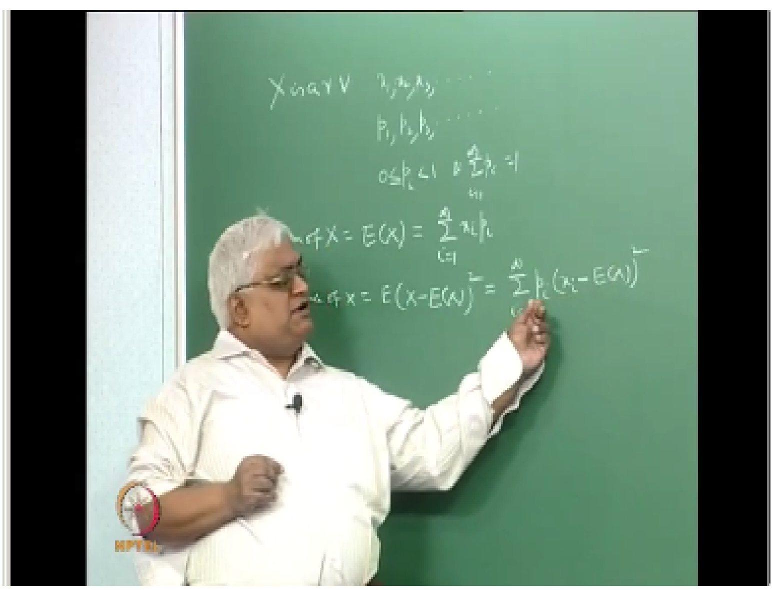
Basics of Statistics, Covariance, and their Properties

Prof. C.A.Murthy

Machine Intelligence Unit,
Indian Statiscal Insistute. Kolkata

In this lecture I shall be talking about basically co variances and some of their properties but then before that in order to introduce co variances I will start with some of the basics of statistics okay.

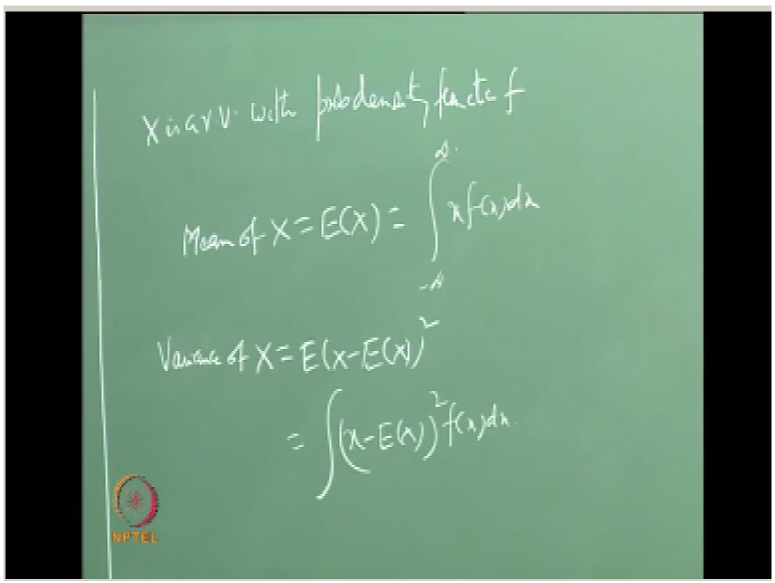
(Refer Slide Time: 00:29)



Suppose X is a random variable taking values let us just say X 1 to X 1 X 2 X3 etcetera infinitely many with probabilities p 1 p 2 p 3 etcetera we are naturally $0 < p$ $\sum$ p I I = 1 to ∞ =1 then the mean of random variable mean of X which is denoted by expected value of x and this is equal to summation i=1 to ∞ x i' x PA this I suppose all of you know it okay then the variance of x which is expected value of this, this is equal to.

So basically ∑ I = 1 to ∞ p I x X i- here please look at it you should write here a value of x a value of x is x I- the mean - the mean the whole square the whole square expectation means this one multiplied by the corresponding probability so this one multiplied by the corresponding probability p I I = 1 to∞ this is for discrete random variables but for continuous random variables.

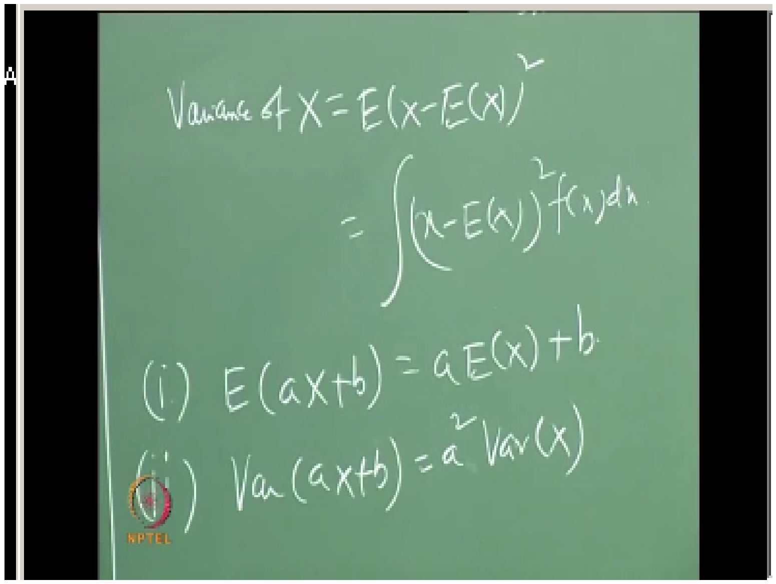
(Refer Slide Time: 02:40)



Here X is a random variable with probability density function probability density function say F. Then me mean of X which is same as expected value of x =∫x FX DX integration over suppose X is a positive random variable that means 4 - ∞ to 0 the density is anyway 0 from 0 to ∞ it is going to be positive okay so even if it is a positive random variable or even if it is not taking some values in this interval - ∞ +∞ one can always write -∞ + ∞ because the place where the random variable is not taking the values.

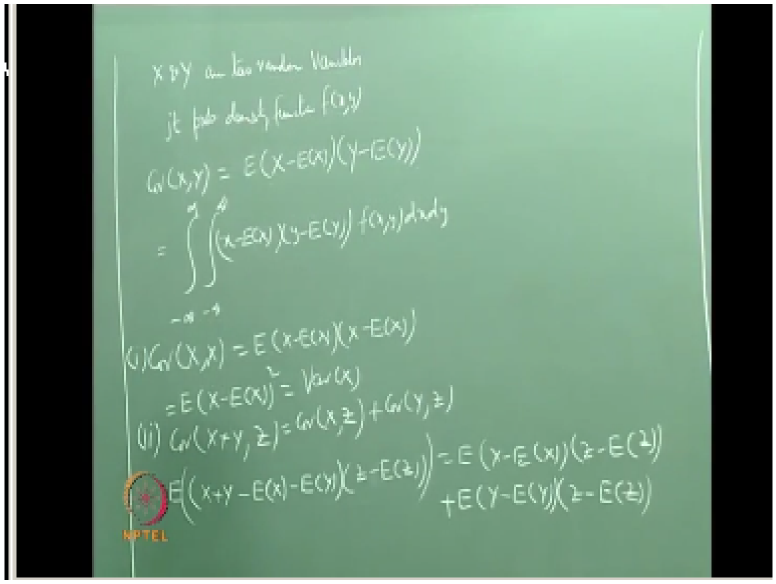
The corresponding density will be anyway 0 so I can always write- ∞ + ∞ then again the variance of X is equal to the formula is the same thing this is equal to ∫ X -okay now there are some nice properties what are the properties first this is whether you have a discrete random variable are you have a continuous random variable the properties that I am going to write they are always true what are the properties.

(Refer Slide Time: 04:46)



Let me write down properties like this expected value of ax + b is equal to this means we have a new random variable Y which I am defining as a times the random variable X + B where a and B are constants so the mean of this random variable is a times the mean of the world random variable X+ B okay now two variants of ax +B is equal to a square variance of X variants of ax + B is equal to a square that means the B does not have any impact whatever B you take does not matter it will be a square variance of X okay.

(Refer Slide Time: 06:36)



Now let us have this covariance so for covariance you need to have two random variables x and y okay and you also need to have a joint probability density function so there are two random variables x and y x and y are two random variables and you have the joint probability density function f x y then the covariance of X Y is equal to basically expected value of this is equal to let me explain note that these small access.

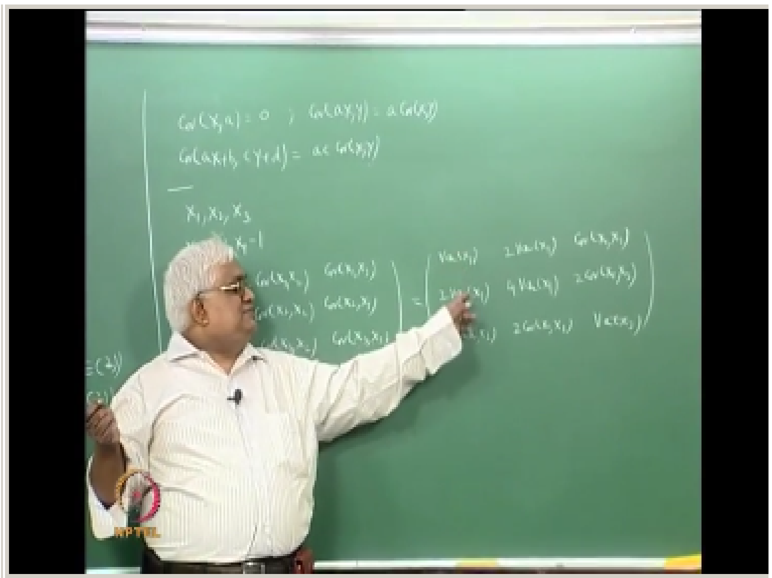
And small wise these are the values taken by the variable and the variables are represented by the capital letters capital X and capital y and this thing expectation of X it is a constant value expectation of Y this is a constant value so covariance of X Y is expected value of x - expectation of X into y - expectation of Y so this is like this the product of X -expectation of X into y -expectation of Y and you should write the density function.

Because expectation you are taking and with respect to X n with respect to Y DX dy and this is go over the whole and you can have similar definition for discrete random variables also which I am NOT writing now well what will be covariance of X with X this will be right and this will be which is nothing but variance of X this is nothing but variance of X.

And there are some more formulas what are the formulas say covariance of let us just say X + y with Z is nothing but covariance of X Z plus covariance of Y Z is this true why just apply this one expected value of x + y - so X + y - expectation of expectation of X + y is expectation of X plus expectation of Y so - is this and multiplied by Z - expectation of Z so this is nothing but expectation of you see X - expectation of X into Jud d - expectation of z.

Then y -expectation of Y x Z -expectation object so X with Z and Y with Z expectation of plus expectation of s so covariance of X + y with Z = covariance of X Z +covariance of Y Z so covariance of X with a constant what is that covariance of X with a constant is zero again you apply the formula X- expectation of X constant is a constant is a the mean is also a -a =0 so multiplication of that thing is zero.

(Refer Slide Time: 13:01)



So you are going to get zero so covariance of X with a is zero and covariance of ax + B with Cy + D this will be first before that covariance of Y X with Y is equal to a times covariance X with y again you apply the formula ax means ax – a into expectation of X so a is going to come out so X - expectation of x into x - y expectation of Y so covariance and a has come out so a times covariance X Y and covariance of ax +B with C y +D = C covariance XY a times C covariance XY okay.

Now suppose we have three variables let us just say we have three variables x1 x2x3 okay and suppose x 2 = two x 1 - 1 x 2 =2.1 x okay now what is now this matrix let us just see covariance x 1 x1 covariance x1 x2 covariance x1 x3covariance x2 x1 covariance x2 x2covariance x2 x3 covariance, covariance x 1 x 1 is variance x1covariance of X 1 with x2 x2 is 2 x 1- 1 so this will be 2 times variancex1 x1 with x1 is variance x1x1 with x2x2 is 2 x 1 - 1 so this with x1 so one times two is two times variance x1and covariance x1 x3.

Let me just write it like this whatever it is they are okay now covariance of X 2 with x1 that is two times right and x2 with X 2 this is 4 variants x1 and x2 with x3 x2 is 2 x 1- 1 this will be 2 covariance and anyway there is a third row now what can you say about the first two rows dependent so what will happen to the determinant zero the determinant value will be 0 so because of just this relationship what is what is happening whatever may be your covariance matrix if two variables are linearly related the rank of the matrix is going to be decreased.

And it will not have full rank it would not have full rank that you can very clearly see and the determinant will become zero okay and the determinant will become zero and I am NOT writing these things this is this will be covariance x1 x3and what will be this, this is two times this will be variants x3 okay it is not really important I am only interested in the first two rows I am only interested in the first two rows.

So if two variables have linear relationship it is going to affect the whole of the covariance matrix I understood it now it is going to affect the whole of the covariance matrix so that as you can see the rank these 2 are linearly dependent rose these two are linearly dependent rose so that the determinant is going to become 0 so you are forced to reduce you are forced to do feature selection in this case to make.

The determinant non zero have you understood what I wanted to say if you get a matrix covariance matrix with determinant as zero okay somehow you should find out which one is doing it or which one is making it so that that particular linear combination or whatever it is somehow you want to remove it so that you will make it the determinant nonzero somehow you have to find the largest minor for which determinant is non zero.

But just minor you need to find out I hope you are understanding the terminology that I am using you understand the word minor you are supposed to find the largest minor and that will actually tell you arm the number of independent variables are it will actually tell you how much rank at most you can make it so you should remove the rest of the things somehow you should find them to be you should find those things.

And you should just remove them because those variables are going to act as noise any question please given this covariance matrix we can say which two variable or related features two features are related look at this one this variance x3 this is somehow not taking any part here so somehow x3 you should take it out you see you look at this two-by-two matrix this row you multiply by 2 you are going to get this look at this with this.

Again this row just multiply it by 2 you will get it look at this with this, this row you multiply by 2 you are going to get this, this sort of thing you would not get if you are looking at this for variance x12covariance x 1 x 32 covariance x1 x3 so you are supposed to get just variance and covariance I mean this one you have non zero determinant probably this, this also has nonzero determinant you are you understanding this nonzero determinant.

This, this is also non zero determination but you have here 0 determination because this is this row if you multiply it by 2 you are going to get this so since these two are these two rows are dependent so here some variable should be removed on the other hand this is nonzero determinant this is also nonzero determinant so here you should keep one variable and from here you should remove one variable.

So naturally the variable that you are going to keep is x3 and the variable that you will be removing is something from here nonlinear related to each other can be using non-linearly later can be used simple no see all these things they are linear have a written anything for nonlinear if it is non-linear have I made any statement about what may be happening no.

So about nonlinear relationship I am NOT I do not want to make any statement on this I am we are only looking at linearly linear case that dependency our independence they are also linearly independent or linearly dependent okay these vectors are said to be linearly independent that is the definition that we read similarly they are said to be linearly dependent if they are not linearly independent.

So I am NOT looking at the non linear part of it and if it is non linear how to go about doing it at least I mean I cannot make any statement using these things that a real serious problem which people are really aware of it but they are not being able to find solutions acceptable to everyone okay they have not been able to find solutions acceptable to everyone I am only talking about linear case yeah.

Now any doubt I can say that you find the largest minor are that largest square matrix with non-zero determine with non zero determinant saying it is easy but then you have to look at all those combinations which may be difficult okay but then the principle I hope you have understood it if your variance covariance matrix has the value determinant as zero determinant value at zero somehow you should find out those variables which are causing this zero.

And you should remove them okay but how are you going to do it that I think the principles are clear to you but then you should act how are you going to implement it which are I mean like you would like to implement it in such a way that it is feasible to implement it I mean you do not want to do an exhaustive search then which will make your life miserable that I think one should try to find out some methods for this and it is not I mean there are some things I think they are already existing in the literature.

The moment you have paper covariance matrix has determined 0 okay there are some papers which you can find the literature about how to go about removing those features so that you would like to make it nonzero so this one I think I will stop here you.

End of Module 06 – Lecture 02

Online Video Editing / Post Production
M. Karthikeyan
M. V. Ramachandran
P. Baskar

Camera
G. Ramesh
K. Athaullah
K. R. Mahendrababu
K. Vidhya
S. Pradeepa
D. Sabapathi
Soju Francis
S. Subash
Selvam
Sridharan

Studio Assistants
Linuselvan
Krishnakumar
A. Saravanan

Additional Post – Production
Kannan Krishnamurty & Team

Animations

Dvijavanthi

NPTEL Web & Faculty Assistance Team
Allen Jacob Dinesh
Ashok Kumar
Banu. P
Deepa Venkatraman
Dinesh Babu. K.M
Karthick. B
Karthikeyan. A
Lavanya. K
Manikandan. A
Manikandasivam. G
Nandakumar. L
Prasanna Kumar. G
Pradeep Valan. G
Rekha. C
Salomi. J
Santosh Kumar Singh. P
Saravanakumar. P
Saravanakumar. R
Satishkumar. G
Senthilmurugan. K
Shobana. S
Sivakumar. S
Soundhar Raja Pandian. R
Suman Dominic. J
Udayakumar. C
Vijaya. K.R
Vijayalakshmi
Vinolin Antony Joans

Administrative Assistant
K.S. Janakiraman

Principal Project Officer
Usha Nagarajan

Video Producers
K.R. Ravindranath
Kannan Krishnamurty

IIT Madras Production

Funded By
Department of Higher Education

Ministry of Human Resource Development
Government of India

www.nptel.ac.in