Data Analysis and Decision Making 3
Professor Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur
Lecture 56

Welcome back my dear friends and dear students and good morning, good afternoon, good evening to all of you wherever you are in part of this globe. And this is the DADM 3 course which is Data Analysis and Decision Making 3 course under NPTEL MOOC series. And this total course duration as you know for 12 weeks. The total contact number of hours for the 12 week course 30 hour which when split into number of lectures is 60, considering that each week you have 5 lectures each lecture being for half an hour.

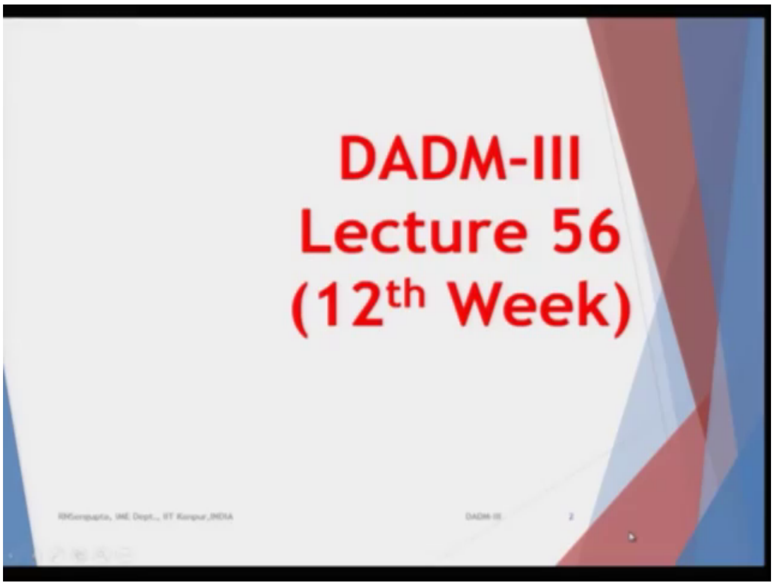
And you have already completed 11 weeks that means after each week you take an assignment so you have already completed by the time you are hearing this. You have already taken 11 assignments that after the 12th one the whole course will end and then you will be taking the final examination. And my good name is Raghu Nandan Sengupta from the IME department at IIT Kanpur.

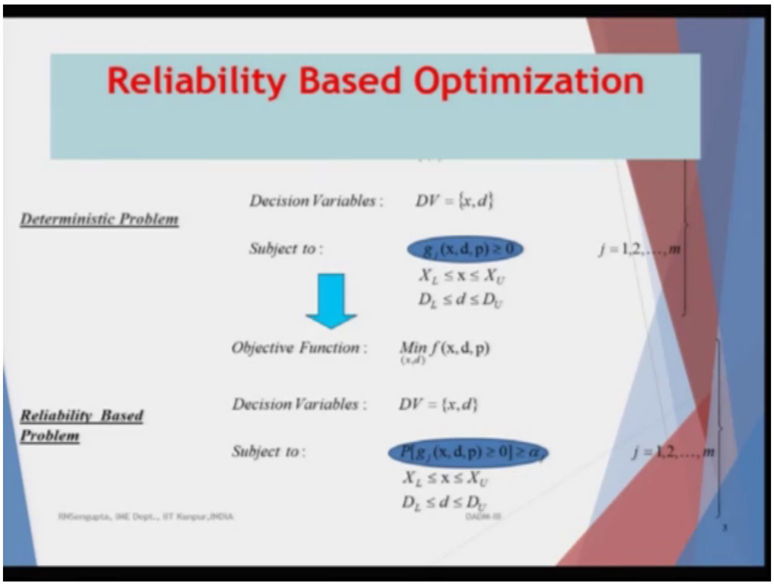
So, in the last two lectures which was the 54th and 5th I just drew many different types diagrams in order to make you understand that considering that normality holds for the distribution whether for the univariate case or for the multivariate case. It was easy to understand how the reliable surface would be the overall area and the center of gravity of the center of the circle or for the ellipse with vertical area the direction being from the major axis or the horizontal direction being the minor axis.

You can find out how the value of alpha as it changes how the locai or each and every point of locus of the reliable solution changes and based on that you can find out that what is your overall the area which should be equal to alpha I's depending on the I as you change the level of reliability on based on that you can find out the solution. And which is the best fit solution for case when the constraints are reliable with certain level of probability.

And for the case when it is non-normal I have mention that time and again that are trying to find out the feasible region is by itself difficult. And also if the center of gravity or the mid or the center of that overall area is to be found out then obviously you will use different types of stimulation techniques and one was the de-couple method, one is the single loop method, one is the double loop method. But I will only concentrate on the decoupled method and consider that how this can problem be tackled.

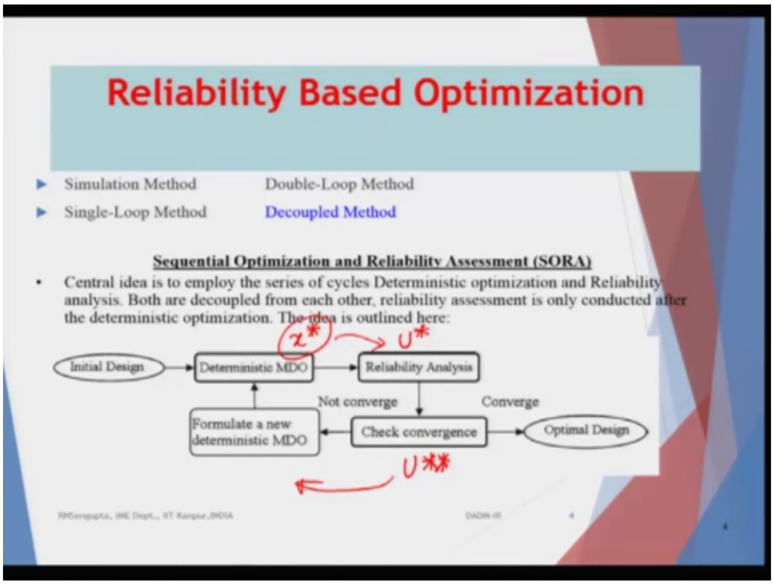
(Refer Slide Time: 02:51)





So this is the 12th week which is the 56 lecture. So here I will again repeat you have the constraint Gjs and j is equal to 1 to m each of them is converted into probabilistic frame work the constraints and this alpha suffix j, j is equal to 1 to m are the constraints and how they can be corporate is what the diagrams I was basically trying to draw.

(Refer Slide Time: 03:16)



So we will consider the sequential optimization and reliability assessment method which is a SORA method. So the central idea is to employ a series of cycles different loops of trying to optimize. The deterministic optimization in one end as I have been mentioning time and again in the x plane or the x space and the reliability analysis to be done in the U space. The transformation basically happens, when you basically use one of the transformation techniques which is the Rosenblatt one.

Which we will consider I will just draw the, write equations without going to the details where you convert from the x space find out x star suffix 1, 2, 3, 4 depending on the stages of optimization that x one star is mapped on to the U space then you find out the optimization the MMP point in the U space again transform it back to x space and continuing doing this.

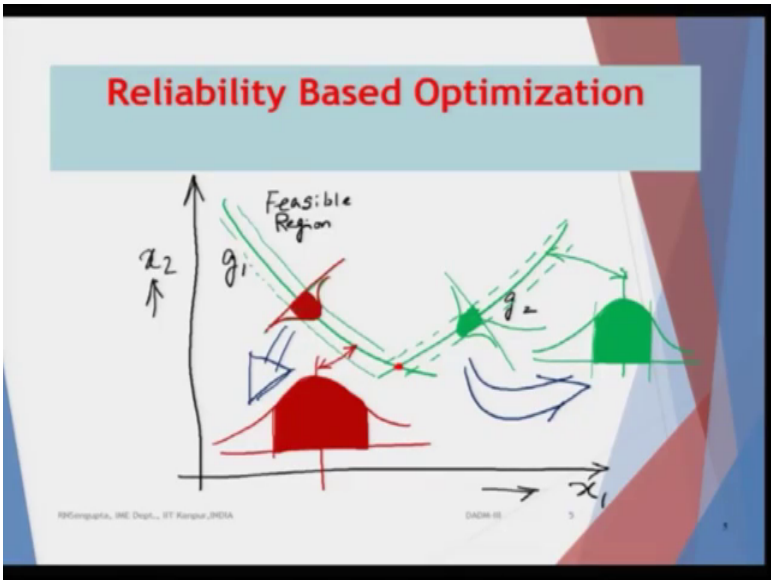
Now this deterministic optimization which you do all the deterministic optimization methods we know whatever I have been discussed. Obviously, I was not able to consider all the mythologies, but whatever you do would be done using the simple optimization techniques we all know or which can be buildup by reading the books. And what is more important is this reliability phase which are decoupled. So that is why it is known as the decoupled method with where you decoupled this deterministic one and the reliability analysis one.

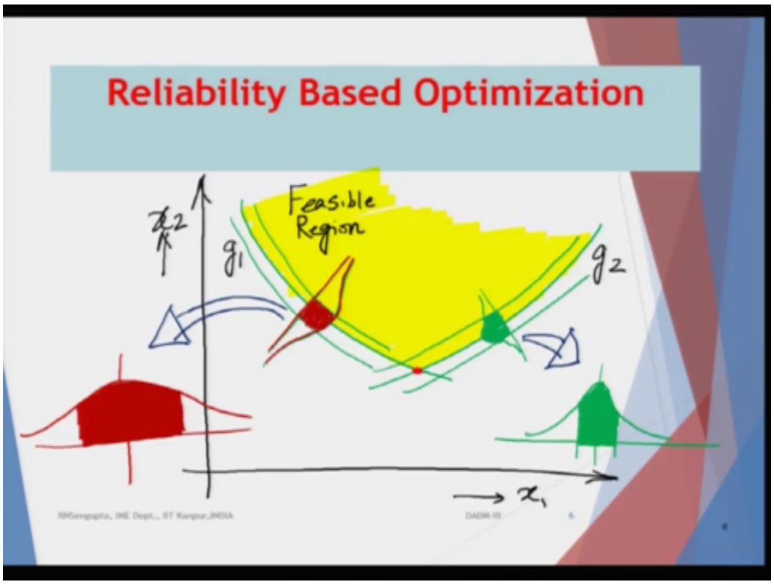
So, let me continue reading it both are decoupled from each other, the reliability assessment is only conducted after the deterministic optimization is done and the idea is like this. You basically have initial design I will come to flow process chart later on. So, generally the bullet points how you do the mythology, you basically design it, design the problem. Design the optimization one and you basically solve the deterministic optimization techniques as such. Once, the deterministic optimization gives you some x star.

Now this x stars need not be scalar they can be vectors also. Once this x star is found out you basically do the reliability analysis and find out the U stars check the conversions if it is not again you basically formulate the new deterministic optimization by basically. So this will be U star and U double star, once U star is optimized U double star. Find out the inverse transformation of U double star into the deterministic space of the x space. Again optimize find out x star the next stage, then again go back to the U space and continue doing this.

In case if conversions is obtained you report those values of U star and those values of x star and your answer or your job is done. So, how you do the reliability analysis is most important which will consider now.

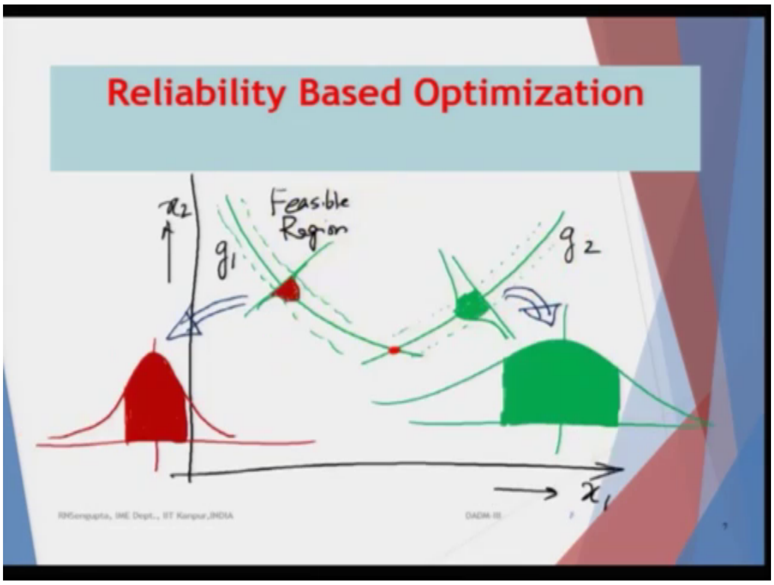
(Refer Slide Time: 06:27)

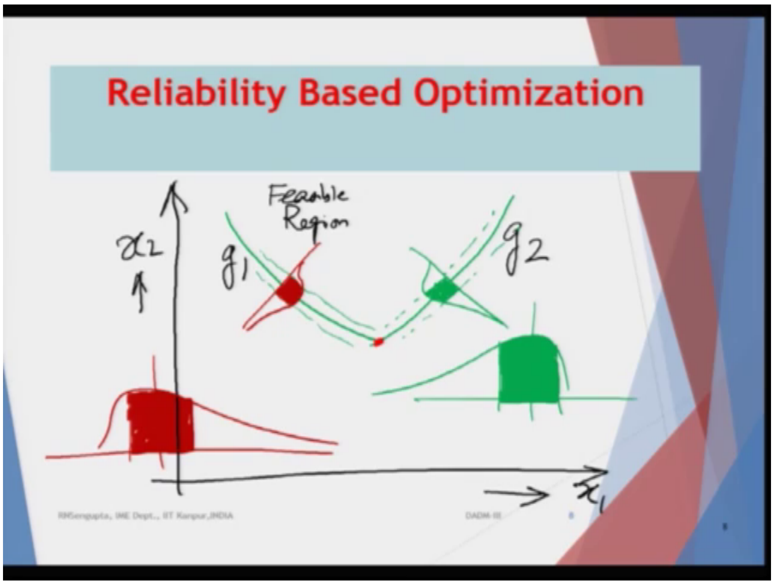




So, these are the diagrams which you have already discussed. And, they would be consider later on when we do the reliability analysis. This is for the case when the variances are same normal case. This is the case when the normal distribution is 2 for both of them, I am drawing only in the 2-dimension one where the variants of x2 or in the y plane is more.

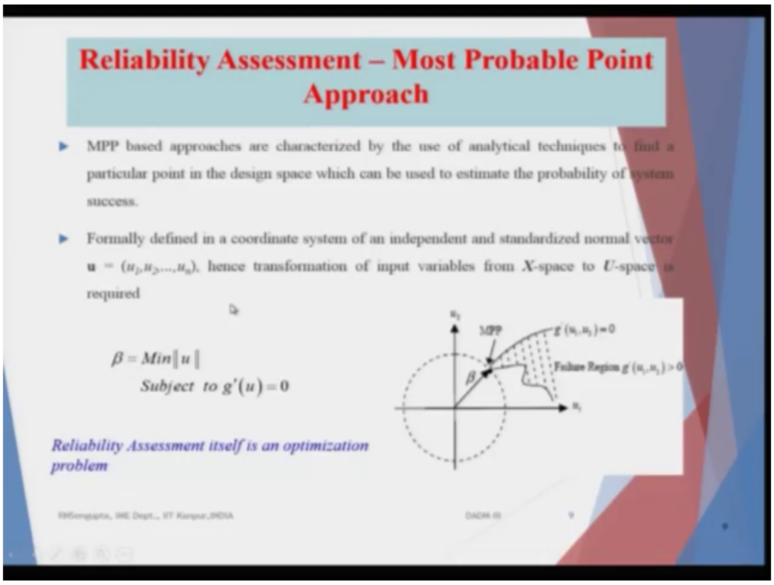
(Refer Slide Time: 06:55)





This is the third diagram where the variance for the variable which is in the x direction and x1 is more. And in this diagram where we consider the last diagram which we also consider in the last class 55th lecture where the distribution per say are non-normal, so I am trying to find out the best surface area of the oral region would be difficult.

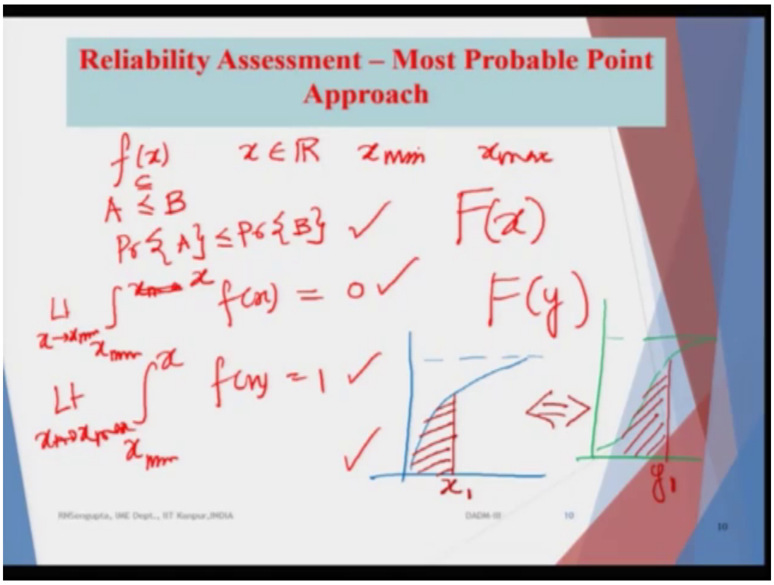
(Refer Slide Time: 07:23)



Now, let me come to the reliability assessment of most probable point approach or the methodology. So, the MPP method is based on and is characterized by the use of analytical techniques solution techniques, to find a particular point in the design space which can be used to estimate the probability of a system and the success. And then you will utilizes that value of U star or U star 1, U star 2 whatever it is in order to report that optimization is being done on the U space.

The formally, define is basically coordinates system of independent and standardized normal vectors. Now, here it should be important that the U space all the variables have been transformed in to U1 to Un and all of them are orthogonal. So why we are doing a let me consider this using the simple Rosenblatt transformation. So, will basically mark it here or try I would not be going in to this diagrams immediately but I will come and mark it accordingly. So, let me prepare a blank slide so it is easy for me to draw.

(Refer Slide Time: 08:58)



So here I start off, okay now the background goes like this you have any distribution for axis. Anyone to transform in to standardized normal space which is U. Now, what is background for that? So the background is we come to simple CDF or probabilities distribution which is the cumulate distribution function. So the story goes like this, consider you have f of x univariate one we will go in the multivariate case within few minutes.

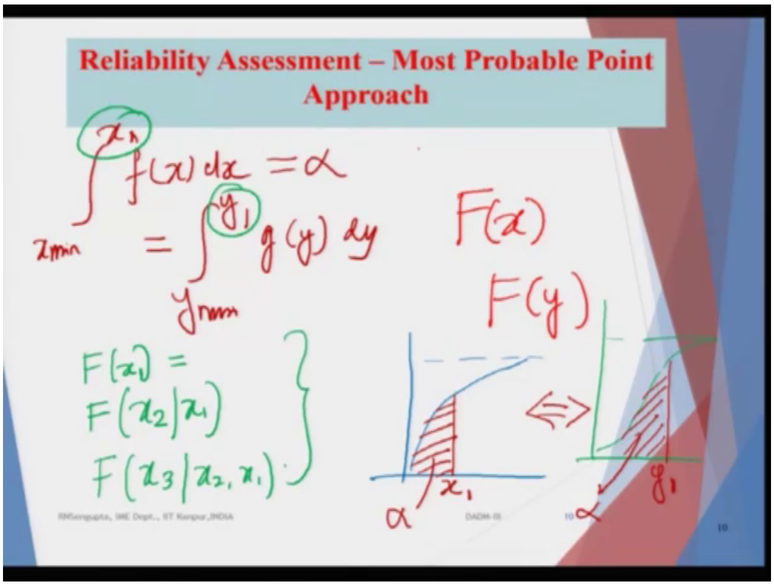
And x is on the real name, so basically x is from minimum, some minimum to some maximum which means if and consider this is the continues one it will be the same, the concept would be the same if we consider the discrete case of x also. So, what it means? All the 4 properties for the CDF should hold, 4 properties if you remember I have done in DADM 1 is that, the if A is less than equal to B, A and B are the events which consist of some of the values of axis.

Then the probability of A or consider it is as a subset it means this is a mono technically increasing function. So, because probabilities cannot be 0, they can be in the sense they cannot be negative. If they are 0 or positive so addition values of x is realized value take the cdf value would be increasing. Point 1, number 2 is that from x min to x max limit I would not use the word x max it should be basically x only (sorry). Limit x tending to x min f of x would be 0, limit x min to x and x tending to x max f of x is equal to 1.

And the other one is that the concept of left continuity and right continuity holds. So, the first one, second one, third one are the left and the left continuity, I continuity concepts. If holds will considered that to be a CDF function. Now what is the beauty based on which we will proceed is this. The minimum value for the CDF is always 0 whatever the PDF is or whatever the PMF is, the maximum values is 1.

So, if you have 2 particular random variables f of x and f of y. We are going to transform and y is basically U here which will consider later on. The distributions diagrams would look like this the CDF values. So, 1 is between 0 and 1 and consider the other one is also between 0 and 1. So I am going to do, one to one transformation between these two. Which means if I know x and I know this CDF value I also try to find of a particular of the same CDF value here. Such that the value of x1 which I have would be mapped on to y1 provided this is true and what is true? Let me mark it here.

(Refer Slide Time: 13:24)



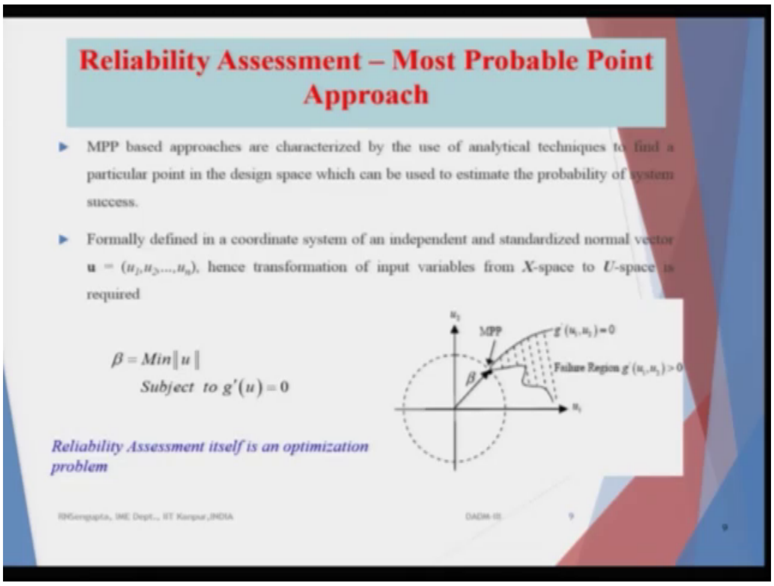
It is like this x min whatever it is till x1 f of x dx will be equal to now should be equal to because this is equal to alpha, this is also should be equal to alpha, this is alpha the overall area is equal to y min because y1. So consider this PDF is g of y dy. So we can say convincingly the values of x1 and the values of y1 should be such that the overall area covered by addition of all the properties from the minimum values to those corresponding values of x1 or ay1 in whichever space they are should be exact equal.

Now, this is what we are going to utilize in the Rosenblatt transformation considering it is multivariate distribution. What it does is that it fixes f of x at some value which is given. And then correspondingly so this is f of x1, x1 is basically now a random variable of 1-dimension what we want to find out is the value of x2 given x1 is fixed we calculate the values and find it out. Given x1 and x2 we try to find out x3 and proceed in this direction such that given the prior point, given the prior values we know that what value would be fixed for x1, x2, x3.

Such that the transformation would be given that at a fixed value of x1, what is the value of x2? And, we are purposely trying to take the orthogonal values of such that no dependence structure is between x1 and x2. Given x1, x2 as orthogonal we fixed them then try to find out x3 value such that x1, x2 are already orthogonal. Then we will basically try to find out the orthogonal transformation of x3. So here when you doing the orthogonal transformation you are basically doing in the U space which as said is the standard normal deviate which are all orthogonal U1 to Un.

Then given x1, x2, x3 are all orthogonal or consider U1,U2, U3 are all orthogonal you try find out U4. And proceed in this directions such that you are able to find out all the values of U, so this is what it states.

(Refer Slide Time: 16:12)



Formally, define it is a coordinates system where they are independent this is important independent and standardized normal vectors. Normal vectors because, all of the mean values is 0 and standard division is 1. Such that the cumulative distribution values which you take for the multivariate cases exactly equal to cumulative distribution in the value terms the alpha value on the x space.

So, let me again read it, so this is a coordinate systems of an independent and standardized normal vector U, U is the vector from U1 to Un, hence transformation of the input variables from the x space to the U space is required and is possible. So what we do is this, we basically try to follow two procedures. We take the value of beta, beta or alpha whatever I have been talking about and we want basically try to minimize the orthogonal the norm or the distance.

Subject to the case that if you remember we have taken G of xdp as less than equal to 0 and that value is given as alpha which is actually beta we are considering here. So we will try minimize the distance covered or the norm of the in the U space. So, U space are orthogonal unit vectors. So if I consider in the 2-dimension case it will be circle in the 3 dimensional case you basically be as sphere and hypo sphere in the higher dimensions.

So, based on the subject that two constraints what I will do is like this and this are will be discussed in a diagram later on. First we will basically try to find out the minimum of the use corresponding to the fact that the first derivative of the use for each and every use which you have is 0. Which means there can be two ways, one is have the boundary space and try to increase the circle like try to increase the circle such that the moment it touches the feasible region n points or the curve or one of the constraints.

That will give me the minimum value of beta, beta is basically the minimum distance remember, so beta can changed such that the tangent point would give you the U star based on which we will basically proceed to do the optimization in the reliability space that is method 1. Method 2 is that keep circle fix at a level of beta and try to shift the boundary such that the boundary the moment it touches.

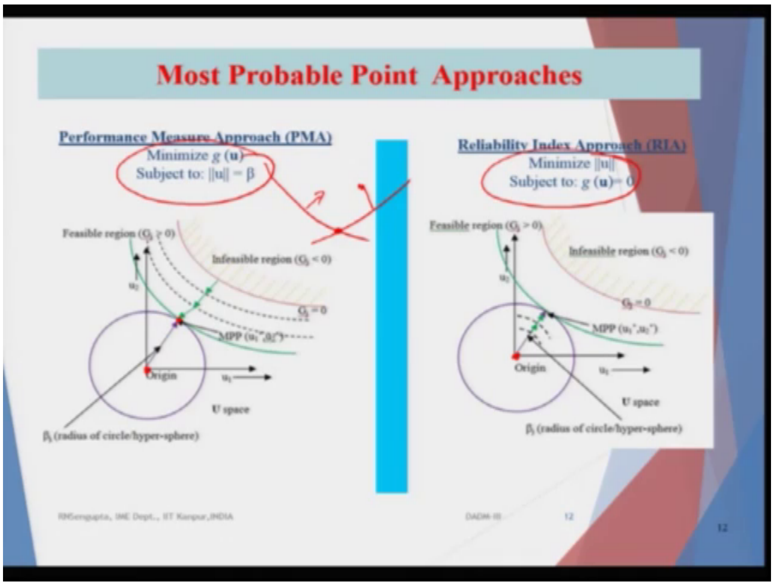
So, obviously the feasible or infeasible regions point would be considered in such a way that the moment the boundary touches the boundary moving. The moment it touches the circle of the sphere or the hypo sphere that point which is tangent will give you the U star. So, this is basically the general procedure. The point where it touches the boundary the circle or the sphere or the hydrosphere that will be the most probable point where the concept of the level of reliability would be maintained and also it will give you the minimum point where we are able to achieve that such that, that is the optimum point in the U space.

Which can be transformed back to the x space to basically start of the next stage of the optimization that means first you do with the minimum values if you remember mu suffix x mu suffix p that value only 1 is obtain it has it will basically proceed to next level which we obtained for the U space again do the optimization based on the second set of iterations in the deterministic space again do the Rosenblatt transformation as I have considered and continue doing it in this way such that you basically achieve the optimality as discussed.

So, now there are the two methods I discussed, one the circler radius increasing or the spheres radius increasing or the hyper spheres over all the volume increasing. Till it touches the feasible region at a particular point and the tangent point at U star or else keep this cycle of the circle of the sphere or the hypo sphere fixed and move the boundary of the feasible region till it is tangent.

So how this are achieved I will just show it very simply in a diagram, I have talked now basically see the diagrams. I will show both the diagrams at the same time rather than showing one at a time and then discuss. So, now here it like looks.

(Refer Slide Time: 20:55)



Remember it is a I am considering it is a circle 2-dimensional variants are same. The first case, you basically have in the U space so you have basically transformed it in the U space you have the origin so the origin is here. And this circle if you remember beta was the minimum of mode of the U space which is the distance. You have basically you want to minimize that, what you do is that you keep shifting the feasible region more inside. So, the area which was there I will just draw it if possible.

If I have it here I will just note it down, yes so this area was there this was the deterministic one. So, you are going inside more inside the feasible region. So, this invisible region is basically now increasing that means the feasible region is now shrinking. You start moving so these green arrows which are there they U they give U the incremental step in which the invisible region increases. And, it goes more inside the feasible region that means the feasible region shrinks till the point is reached where it is tangent.

And that tangent point which I have mentioned is the MPP or the most probable point in the two dimensions case which would be U1 star and U2 star in the first iteration. So, it gives you the infeasible region for the case when you have Gj which is the transform function which you have in the U space. And, the feasible region now shrinks it moves step by step and these dotted lines which you see, you start with the red once the dotted ones are as they expand or the feasible region decreases till it is tangent.

And the points which you have the coordinate system which you have in the two dimension 1 is U1 and U2 which are always orthogonal as I mention. So, this is the method which you have for the PMM method which is the performance measure approach where your main criteria is to minimize the functions g of u provided the mod of the distance for in the U space is beta, beta is the level of probability which you have already define by the decision maker.

While in the reliability index approach which is RIA, you just reverse, reverts in the sense you minimize the mod of the distance subject to the case that g of U or the functional transformation which you have done for G in the U space is kept at 0. Because 0 is the basically if you remember is he boundary. In this case it happens like this you have the infeasible region and feasible region is demarcated very clearly.

Now, the circle with a point which is U1, U2 or the origin whatever it is there, it is starts increasing like concentric circles. The moment the value of the minimization is approached subject to the case when you have the region when it is exactly equal to 0. Because, the demarcation is that in the feasible region Gj, Gj for each constraints is greater than 0 in the feasible region is less than 0. So at the boundary it will be exactly equal to 0.

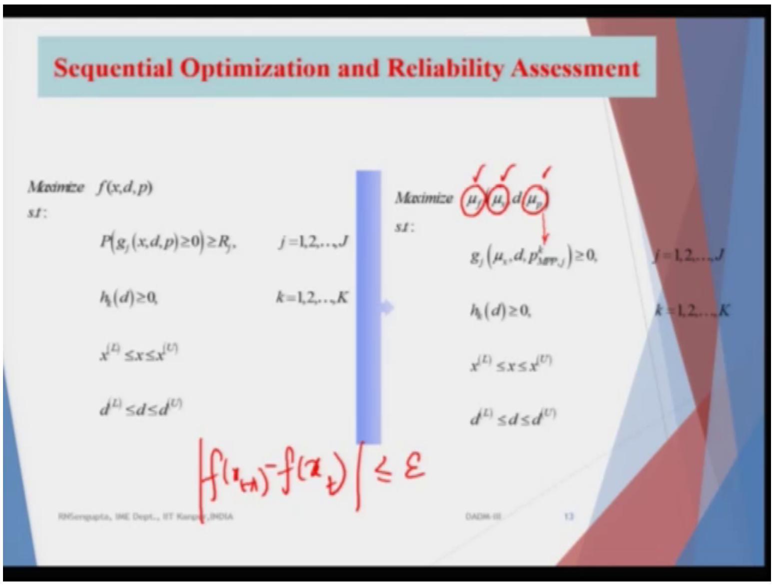
The moment it basically increases and becomes a tangent at that point. The radius of that circle which will be beta which we are trying to minimize. So, minimization of the mode is basically radius that will give the level of reliability which, we want to achieve based on the fact what has been given a priory. And, the point where it will be minimum for the optimization case which is minimization of the mod would give you that U star which is the MPP point in U1 star and U2 star.

See, one case I am repeating the boundary moves inside that mean the more inside the feasible region moment it becomes tangent, you stop, report that values and that point is the MPP which would again transform back into the x space and we will start the second iteration using the deterministic one. So, the first iteration was I am again repeating it, based on the mean values of the x's and the P's and then the second iteration will start, you will get a answer you will again do the resentment transformation come to the U space and again do either the PMA or the RIA approach.

RIA approach is again I am repeating, the circle starts expanding the movement it is tangent you report that U star values U1 star, U2 star and that radius is basically the level of beta. So, if you look at this optimization, they are not like the primal dual problem but, you will are basically trying to convert the subject to constraints as an optimization minimization problem from the PME to the RIA.

While the minimization function is transform into subject to constraints in the RIA which is g u is equal to 0 because the 0 point is the boundary where is the feasible region the infeasible are demarcated.

(Refer Slide Time: 26:37)



This concept which you have already discuss but I will again discuss in more details. So your optimization problems in the case of the probabilistic framework is like this, maximize or minimize whatever it is. So, remember this maximization minimization I am talking about in the x space not in the U or the reliability one which will always be a minimization one because we want to find the most probable point.

So, the maximization is done for f of x which when transformed in to the reliability one is basically the mean values of the function and the mean values for the probability for the parameters which are probabilistic. So, which are this I will just mark it and obviously for x, so there are 3 minimum values where we start. And based on that you when you do it you find out the most probable point for the p which is the probabilistic point or it can be for the x also the depending how the problem formation is been done.

And your repeat from the x to the U and U to the x continuously till you are able to achieve this. So this is for the Tth iteration and this is for the Tth plus 1 iteration, the mode of that should if it is less than or epsilon you stop it. So I will discuss few problems and the concept later on in the 57^{th}, 58^{th}, 59^{th} and 60^{th} lecture. And in the 60^{th} lecture it would be wrap one and I will continue discussing about the SORA method, the PMA method and RI method and also going to robust optimization later on in the last 4 classes. With this I will end this lecture and have a nice day and thank you very much.