Data Analysis and Decision Making-III
Professor Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur
Lecture No. 38

Welcome back my dear friends, a very good morning, good afternoon and good evening to all of you wherever you are in this part of the globe and this is the DADM III which is Data Analysis and Decision Making 3 course under the NPTEL MOOC series and as you know this is the third part after DADM I where we considered statistics, DADM II whatever basically all the non-parametric methods and this is optimization and operation research.

And this total course duration is for 12 weeks which when converted basically is thirty hours contact hours and which is basically 60 lectures and each week we have 5 lectures each week for half an hour and after each week we have assignments and as you can see you are in the 38th week and which means you have already taken 7 assignments and you will be (taut) totality taking 12 assignments after that there would be the final examination.

My good name is Raghu Nandan Sengupta from IME department at IIT Kanpur. So we were discussing about the optimization problem where your maximization concept was basically trying to find out the efficiency in the numerator you had the bundle of the outputs and in the (numer) denominator you have the bundle for the inputs for if they were the capital K number of DMUs where capital K is basically 10 if I consider.

I have such ten optimization problem maximization of these ratios and the constraints of for all these 10 optimization problems are exactly the same where you want to basically ensure that the efficiency for the first second third fourth till the tenth one where the efficiency are by definitions bundle of outputs for the first one divided by the bundle of the inputs for the first one which is the efficiency of first one similarly if the efficiency of second one third, fourth each being less than equal to 1.

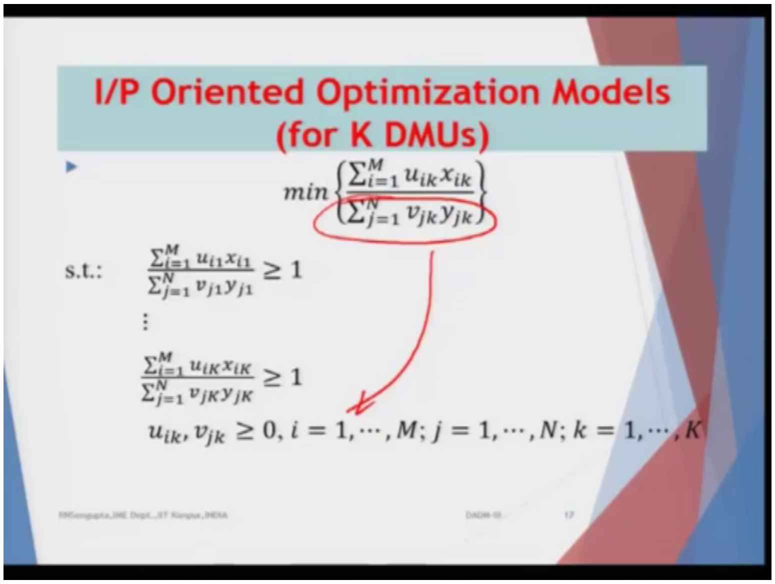
As this is the nonlinear (pro) optimization problem we convert it into a very simple linear programming by restricting the denominator which is the bundle of inputs as 1 so it is basically forced in to the constraint so and each of the other 10 constraints considering capital K is 10 they are now converted in to in inequality sign that means I take the denominator on to the right hand side again to the left hand side so now you have bundle of outputs minus bundle of inputs less than equal to 0. So 10 of them the which is capital K number of them and the eleventh one is basically the corresponding bundle of inputs for the first optimization problem similarly the eleventh constraint for the second optimization problem is corresponding to the bundle of inputs for the second (op) DMU so and henceforth till the last one so now you have uh simple linear programming problem which is objective function is linear constrains are linear and you solve it accordingly.

Now let me come to the input oriented model where the idea which you have got in the output oriented model will just be reversed in the sense now rather than maximizing the the ratio of the output to the input and now basically minimize the ratio of the inverse which it means minimizing the ratio of the input to the output where we are trying to basically minimize the 11 by the efficiency so minimizing that was basically means maximizing the efficiency also.

So in the minimization problem I have in the numerator the bundle of inputs in the denominator I have the bundle of outputs so if there are K number of such DMUs K is equal to 10 as an example for this case also you will have such K objective functions minimization of them the constraints are exactly the same but in this case remember the constraints are greater than (equ) equal to 1 because now the ratio as if you can see the ratio is basically the bundle of inputs in the numerator and bundle of outputs in the denominator.

So basically we will consider bundle of inputs for the first one divided by the bundle of outputs for the first one is greater than 1 then the second constraint for the (fa) first objective for the first optimization problem would be the bundle of inputs divided by the bundle of outputs for the second one is greater than 1 similarly the third constraint is bundle of inputs of the third one divided by the bundle of outputs for the third one is greater than or equal to 1 so and so henceforth such 10 constraints and it would be repeated for the first (op) optimization problem second, third, fourth till the tenth one.

(Refer Slide Time: 4:46)



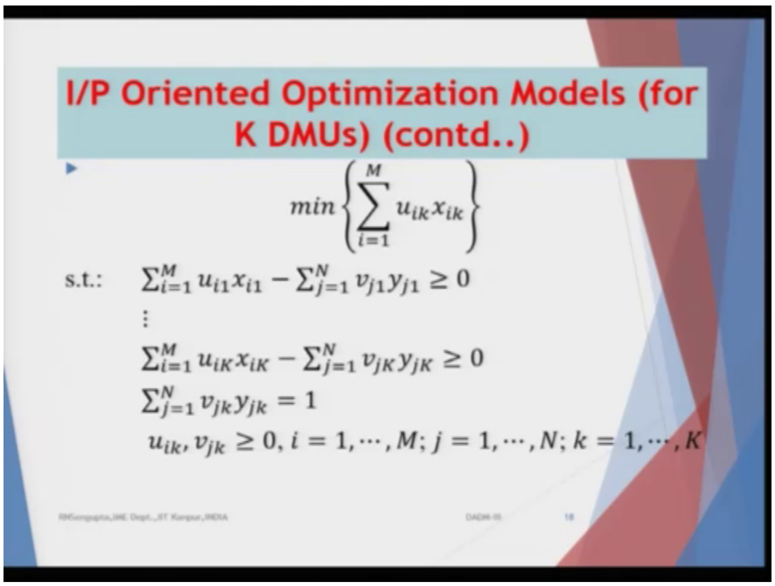
Now again by the same logic as they are non nonlinear we convert into a linear one what we do is that we take the bundle of outputs which is there, force it to 1 that means normalize it to 1 and bring it in to the the constraint so which I will show it within few about 1 minute so now you will basically have a minimization of the of the bundle of inputs which is a linear programming and the constraints are as they were greater than signs, so I will what I will do is that I will take the op outputs on to the right hand side then on to the left hand side.

So obviously they would be converted into the problem where the input minus the output is now greater than equal to 0 so each (obj) see each optimization problem is now converted in to, I will just repeat the first, second and in the similar way it will go for till Kth one which is the tenth one so the objective function is minimization of the bundle of inputs for the first one and the constraints are as follows.

Bundle of inputs for the first one minus bundle of in outputs for the first one is greater than 0 second constraint is bundle of inputs for the second one minus bundle of outputs for the second one is greater than 0 so and henceforth for the tenth constraints and the eleventh one would basically be the denominator which is coming from the objective function so in the first (object) optimization problem it would basically be the bundle of outputs is equals to 1.

Similarly the eleventh constraint in the second optimization problem would be the bundle of outputs for the second DMU is equal to 1, similarly for the tenth optimization problem all the 10 constraints remain as they are that means bundle of inputs minus bundle of outputs is greater than 0 and the eleventh one or the capital K plus 1 constraint would basically be the corresponding bundle of outputs for the last (op) optimization problem is equals to 1 and the minimization problem obviously remains the minimization the of the the input bundle so let us see and obviously it is true that u and v are greater than 0 which are the the weights which we are trying to consider the decision variables which are x and y.

(Refer Slide Time: 6:54)

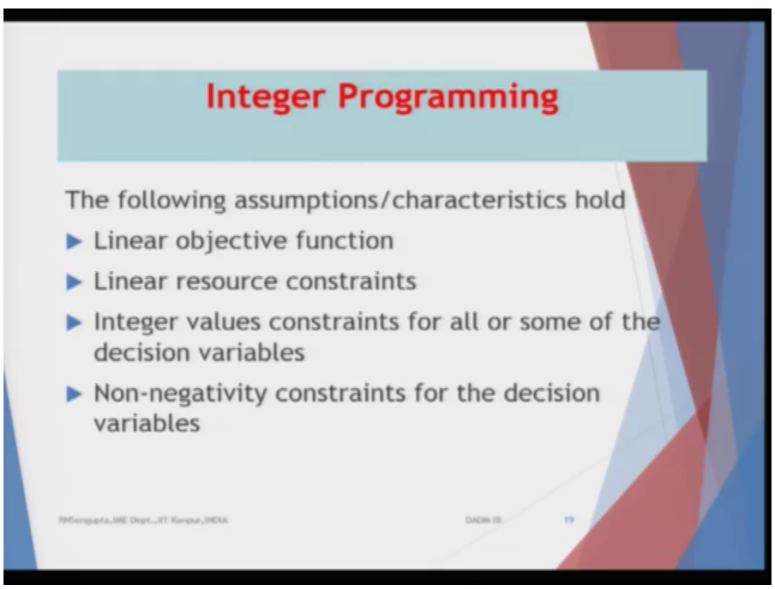


So here you have minimization of the bundle of inputs and the constraints the first K number of them are the bundle of inputs minus bundle of outputs are greater than 0 for the first, second, third, fourth till the tenth one or the Kth one and the eleventh one is basically the corresponding one with respect to the optimization problem which you are doing for the which DMU first, second, third, fourth. So that bundle of outputs will come in to the as the eleventh constraint so there the bundle of outputs equal to 1 because we have normalized them as 1. And then you solve it in the same way as you have done for the linear programming. So here it is the greater than type so in the other case it was less type so you will basically add the slack and surplus and if you require you will add the artificial variables also and will solve it the using the simplex method having the tableau.

And obviously if the number of DMUs of the number of constraints are very large or small you will be can convert it into a dual problem also and then solve it accordingly. Now we will go in to the concept of simple the ideas of integer programming now the integer programming the word says that in this integer programming your main concern is that the decision variables are integer that means the final answer should be integer. Now if you remember when your concept started the concept or linear programming if the problem was to solve say for example number of liters of paints to be produced or number of say for example length of wood to be cut so in that those cases the variables would be safely considered as continuous but obviously greater than 0 continuous means it can be 2.25 liters it can be 3.69 feet whatever it is.

But now if you are if we were considering the number of chairs to be produced number of tables to be produced number of trucks to be transported between destination or the ware houses to the to the retailers or the number of boxes are to be packed so these cannot be continuous they have to be integers like 1 box, 3 (ta) lorries, 4 tables so and so henceforth. So in order to solve the problem obviously we will take the records of the concept of trying to solve them using the linear programming but there would be difference in the how we solve the problem. So I am going to come to that in a little bit detail that how you solve the problem and how they you can get answers accordingly.

(Refer Slide Time: 9:31)



So the following assumptions would be or the characteristics would hold for the integer programming number one the linear objective function concept would hold so that means you have a objective function which is linear in nature like summation of ci into xi or like if you consider the the transportation problems which you have consider, where you have consider the north west quantum method, the Vogel's approximation method all these things.

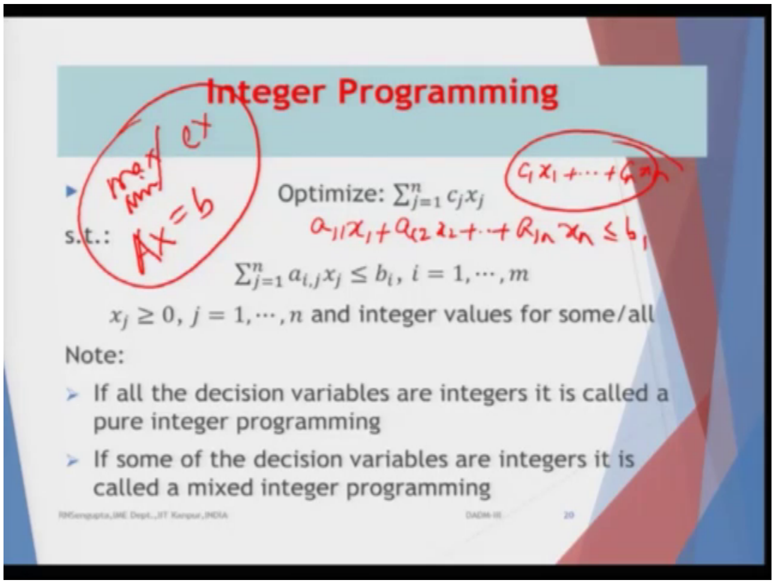
You will also consider linear resource constraints as it should be, so the constraints who are less greater than type, less than type, equal to type and then also you can also add slack, surpass, artificial variables as the case may be as you solve the problem and obviously the concept to solve the linear programs are always the same that means you will consider them to be solved using the concept of this this trying to follow the corners and and use concept of linear programming simplex method.

The integer values constraints for all or some of the decision variables would be true. Now here is is the main cracks of the problem you are considering the constraints for all or some of the decision variables or as it is but the problem is that some of the decision variables or all of the decision variables are integers like if our problem was basically to (manuf) to (ca) cut the wood or or to cut the tree to make chairs and tables and you can cut the any length of of wood or a log.

And consider they can be any (cont) continuous variables in centimeters or millimeters but the number of chairs or tables to be produced are integers only. So obviously it will become a class of problem which is known as a mixed integer linear programming because you will basically have a mixture of some decision variables which are continuous and some decision variables which are integers but if all of them are are integers it is a pure integrally linear programming, linear programming because the objective function is linear.

So the integer value is constraint for all or some of the decision variables will hold true and non-negativity constraints for the decision variables will also hold true because number of chairs cannot be negative number of log which you want to basically cut in order order to basically make the chairs and tables as even if they are continuous cannot be negative so you will consider only a important change which is happening that the decision variables some of them or all of them can be integer.

(Refer Slide Time: 12:21)



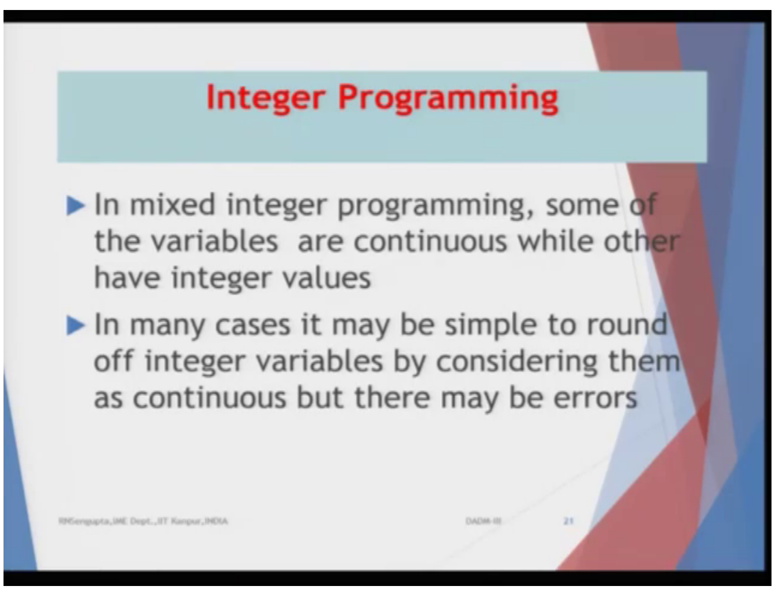
So the problem formulation as I as I said would look like this consider you basically have i is equal to 1 to m that is the number of constraints and j is equal to 1 to n is basically the number of such decision variables you have. See here objective function is maybe say for example a linear one which is a summation of c1 into x1 plus c2 into x2 that means the cost or if you consider the problem that number of such trucks you want to transport on the number of say for example, so cs are the costs and xs are the basically the number number of trucks which you want to apply between roads.

So here I am not considering the the per unit cost for the but the truck trucks which traverse the road so what is the cost I am not going to consider that I am basically, simply formulate in the problem as the number of trucks which as x as x js and cs are the cost which I am considering as to be constant that means the moment I have I hired the truck I have to pay basically a down payment. So I optimize the linear combination of cjs into xjs and also I consider the constraints are less than type so it can be greater than type also it can be equality also so less than type greater than type equality type obviously bring in to the picture the concept was slack surplus on the and if I consider the concept of of this artificial variables so I have the constraints as aij so I have the a metrics capital A metrics.

So these are the constraints which I am I am I am trying to bring to the picture which is aij into xj is less than equal to bi and i is equal to basically 1 to m so I have problem formulation as c1 x1 plus cn xn so this is the objective function I have and the constraints are a so I am going considering the row wise so basically it is a1 a11 x1 plus a12 x2 plus dot dot plus the last one a1last is n xn is less than equal to b1. So if you remember the ax into that when you are solving that concept using the Gauss Jordan concept so this that was to basically maximization or minimization whatever it is of cx so cnx are all vectors and you have ax is equal to b where x is a vector b is a vector and a n is a metrics.

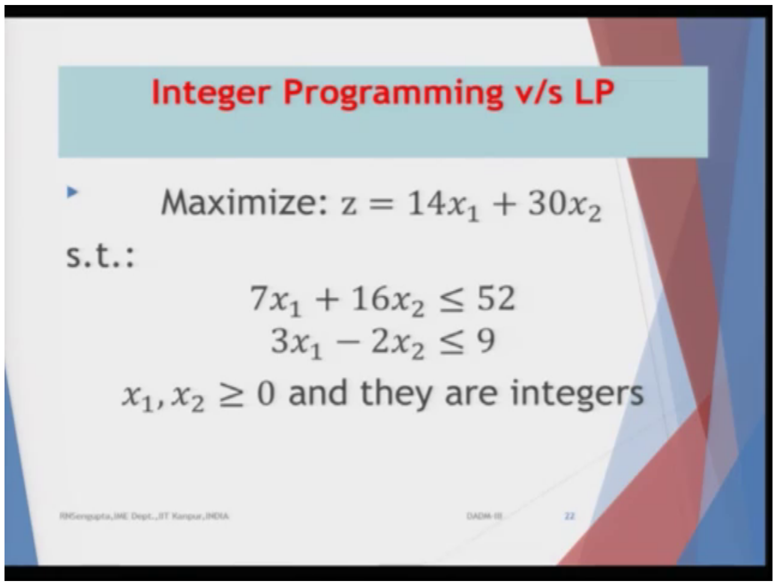
So this problem I have already discussed so I consider that this is the first row then corresponding to the second row it will be the a it will be a21 x1 plus a22 x2 dot dot till a2n xn is equal less than or equal to b2 similarly the last one will be am1x1 plus am2x2 dot dot till amn xn is less than equal to bm. So here xjs are greater than 0 j is equal to 1 to n and integer values for all or sum would be applicable. So note if all the decision variables are integers is called a pure integer programming if some of the decision variables are integers and some are continuous it is called a mixed integer programming MIP. So I will consider the solutions in the very simple concept initially for both of them.

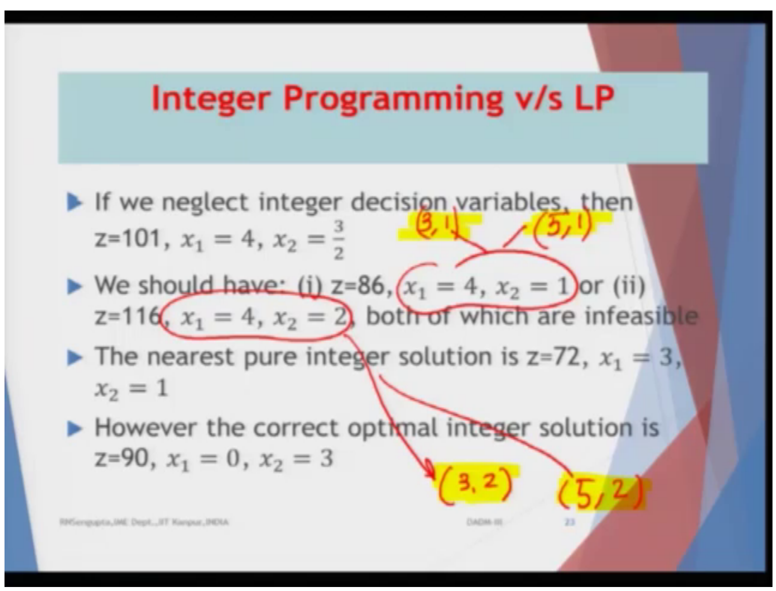
(Refer Slide Time: 16:06)



So in mixed range of programming some of the variables are continuous while others have integer values as already discussed. In many cases it may be simple to round off now here where the concept will start in many of the cases it may be simple to round off the integer values by considering them as continuous but there may be some error, so if say for example the value comes out to be 3.2 I can either take say for example 3 or I can take a value of 4 and try to basically to find out that which that values are possible and obviously remember when I use the word possible that it means that whether this (pro) values of xs or decision variables are also among the feasible points. It may happen the taking the integers values some of the points may be out of the feasible region which is not where solution is not possible.

(Refer Slide Time: 16:57)





So let us consider this problem simply, so your problem is or we want to maximize it can be maximization minimization whatever it is it does not matter and I am considering it two dimension one so z is equal to 14x1 plus 30x2 is basically the maximization problem and the constraints I am considering for our simple case again I am I am repeating it can be less than type greater than type it does not matter. 7x1 plus 16x2 is less than equal to 54 and 3x1 minus 2x2 is less than equal to 9 and x1 and x2 and all of them are integers so obviously now where the concept will start in again I am saying in a simplistic way we will consider so if I solve the problem and if there was nothing was mentioned about integer fact obvious we will get x1 and x2 and they can be integers, they cannot be integers also they can be continuous also so I plug them in a maximization problem I get the answer, I am happy and I solve the problem.

So let us proceed one by one if we neglect the integer decision variables then the values which we will get for x1 and x2 is like this x1 is 4, x2 is 3 by 2 if you put thus problem solution in this problem, so what was the problem, this 14x1 plus 30x2 so it will be 14 into 4 plus 30 into 3 by 2 if I solve the problem my answer is 101 and I am very happy so I have solve the problem. Now what is there see the problem x1 is 4 which is an integer but we also mentioned that x2 should also be integer so it is 3 by 2, so I will be tempted so 3 by 2 is what 1.5 so I will be tempted to basically solve the problem for x1 is equal to 4 x2 is equal to 1 in one case and in another case I will try to I will be tempted to solve the problem of x1 is equal to 4 and x2 is equal to 2 and then get the solution.

But answer may be different which is fine we get a higher value considering the maximization problem but whether they are feasible that is the main point. So we should basically have x1 is equal to 4, x2 x2 is equal to 1, the actual z value is 86 for the first case or if x1 is equal to 4, x2 is equal to 2, the z value is 116 both of which are infeasible. Which means the value of x is equal to 3 by 2 and any value of x2 is equal to 1 or x2 is equal to 2 does not and considering the point of x1 is equal to 4. So the coordinates of I will I will I will repeat the coordinates as the first is for x1 second one is for x2 so a 0.41 or a 0.42 is not inside the feasible region that means they are not contenders to be consider at all when you are trying to solve the problem so they are infeasible points.

The near is now let us consider what is the nearest one. So once we basically have so intendedly what you will do, you will basically change x1 keep as 1 now you will basically change x2 to 3 and another case you will change x2 to x1 sorry my mistake x2 keeping as fixed as 1 you will basically change x1 to 3 and in the next case you will change x12 to 5 similarly you will consider x2 as 2 and then change change x1 in one case as 3 and other case in as 5 that means you will have such four combinations.

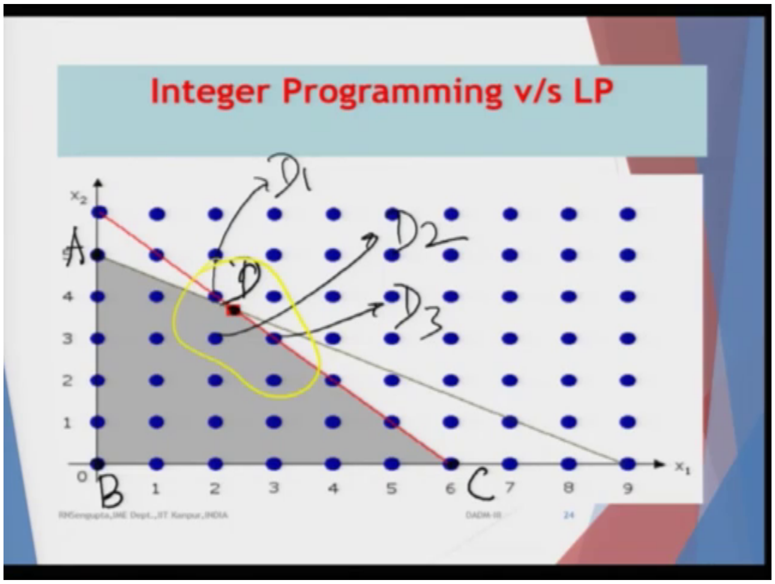
And you will try to basically do it iteratively but the answer may not be right so let us consider one of them so the nearest pure integer solution is when I consider as I mention so initially what I considered was what what I was mentioning is is I if I consider this combination so I had 2 so one point is basically 3, 1 and another point I was going to consider 5, 1 next point would be I would be consider a point 3, 2 another point would be 5, 2. So I keep checking them so one of sum of them will be feasible some of them infeasible some them may be may be inside the feasible point so I would not be able to get the the best optimum solution so the nearest pure integer solution for this case is when x1 is equal to 3, x2 is equal to 1 plug it in this value so the actual uh z value would be 14 into 3 and 1 into basically 30 so in that case you will basically have 42 plus 30 which is 72.

But interestingly the correct optimum solution is no is not any one of them what are these values which I have drawn or written down is not this, not this, not this, not this. What is this let us see. However, the correct optimum integer solution is not this it is basically x1 is equal to 0, x2 is equal to 3 and when you use plug it in the value that means I have to basically go do the search in such a way that rather than changing the (inte) non-integer value in to into integer, which I will do it but first I will be basically more interested in trying to basically fix the non I am just saying it there are methodologies for that.

I will basically fix the the non-integer value into integer one and then basically do the searches for the integer value which was already x1 and they basically try to change it in steps of one then in steps of two and basically do search which is a very rigorous and a brute force but we will have basically have methodologies for that. So, however the correct optimum integers solution is basically x1 is equal to 0, x2 is equal to 3 and the actual z value is 90 which means that in the case of the continuous case you had x1 is equal to 4, x2 is equal to 3 by 2 answer was 101 and in the integer case x1 is equal to 0, x2 is equal to 3 and the answer is 90.

The interesting fact is not the difference in the objective function in this case so what what is interesting is and what I why I am bringing this example for discussion is that, the difference in the in the z values is is I understand it would be 11 but the interesting fact is if you check the x1 value in one case it was 4 in another case it was 0 and x2 value in one case it was 3 by 2 in another case 3 that means the range or or the case of the integer value solutions are such that are trying to find out the integer solutions using the concept of linear programming is is may be as a starting point a good idea but how you basically iterate and find out the optimum point is the main concern which we are going to discuss.

(Refer Slide Time: 24:40)



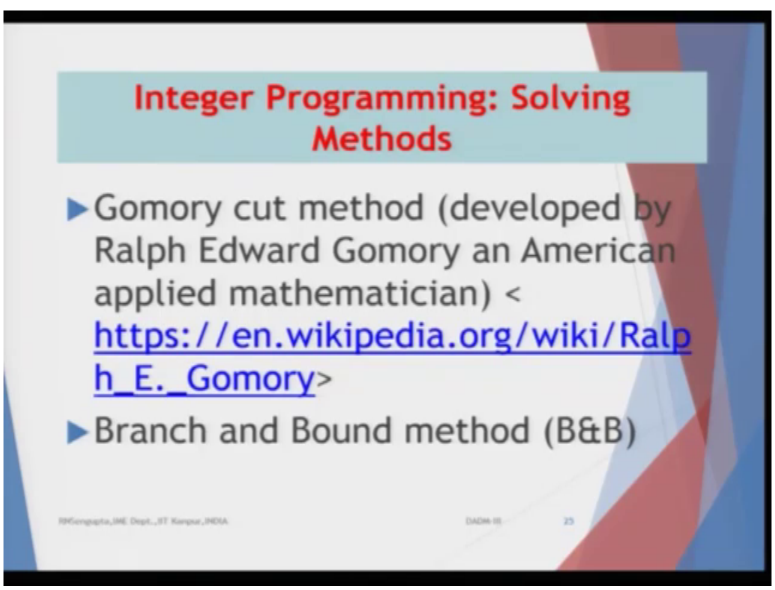
So let us consider this diagram, in a 2 (dimension) dimensional one and obviously in the higher dimension it would be true. See if you consider these are the grid points which I mean drawn so and it is a 2 dimension one it can be extended to the as I just mentioned in the higher dimension so you have the grid point 1,2,3,4 and is going on the right hand side, grid points are also going increasing on the y axis and all these points are the coordinates which are integers.

So any point which is not the blue point in the integer programming problem is not an optimum solution is not feasible obviously now if I have these two lines, I am just drawing the constraints so one of the (cons) constraints lines is the red one I am not writing the equation I am just want to highlight the area and another point is basically the green line which is this one where the pointer (()) (25:24).

So in this case when you solve the problem in the continuous case I will use the color let me use the color black, so the points to solve the corner points is A, B, C, D. So if the problem solution was B, it is integer even though both are 0 either problem solution in the in the linear programming or the case where you consider the continuous one it was A it was a solution in the integer case also. If the linear programming, consider in the continuous case gave the (prob) the solutions C it was integer problem but the interesting fact is D in case if we get D it is an optimum solution because the objective function is moving if we remember is moving outside into more into the first coordinate division now is a maximization problem, it will cross the point D, D is the objective function which give as a point which gives the maximum problem and we stop it.

But the issue is that the moment you try to move the D which is all non-integers, this point which I will consider as D1 is not inside the feasible region so is ruled out. This point D2 is is feasible but is not optimal B is inside the point and luckily in this case D3 is a point which is inside the feasible region and it can be contender of the integer programming problem which we are going to consider. So here the problem which we have considered when you do the solving that means x1 as 4, x2 is equal to 3 by 2, the final answer is x1 is 0 and x2 is 3. So that will give you an answer an idea that are trying to find out an optimum (sol) solution in an around the linear programming solution in the continuous case may seem interesting but trying to find out the optimum result may not be the best way how you do that but the best way in the metrology what we are going to consider.

(Refer Slide Time: 28:04)



So we will consider two methods I will just read it out and then discuss that in the next class. So one is basically the Gomory cut method which was developed by Ralph Edward Gomory an American applied mathematician somebody can have a look at the Wikipedia information about the mathematician and another is basically the Branch and Bound. So in the Gomory cut method what you do is that you shift the boundaries in such a way that you basically eliminate or include some of the constraints in in a way that they consider the integer points also and the Branch and Bound basically I I proceed in in such directions there are there are two paths one if is a maximization problem the minimization problem I proceed in the in the direction where you will basically keep my (()) (28:47) problem in a higher skill.

That means if a maximization problem I take one one of the path where the objective function increases and in the minimization problem it is just the reverse where the objective function decreases depending on the which path I take. So with this I will close this lecture and have a nice day and thank you very much.